DOCUMENT RESUME

ED 370 774 SE 054 331

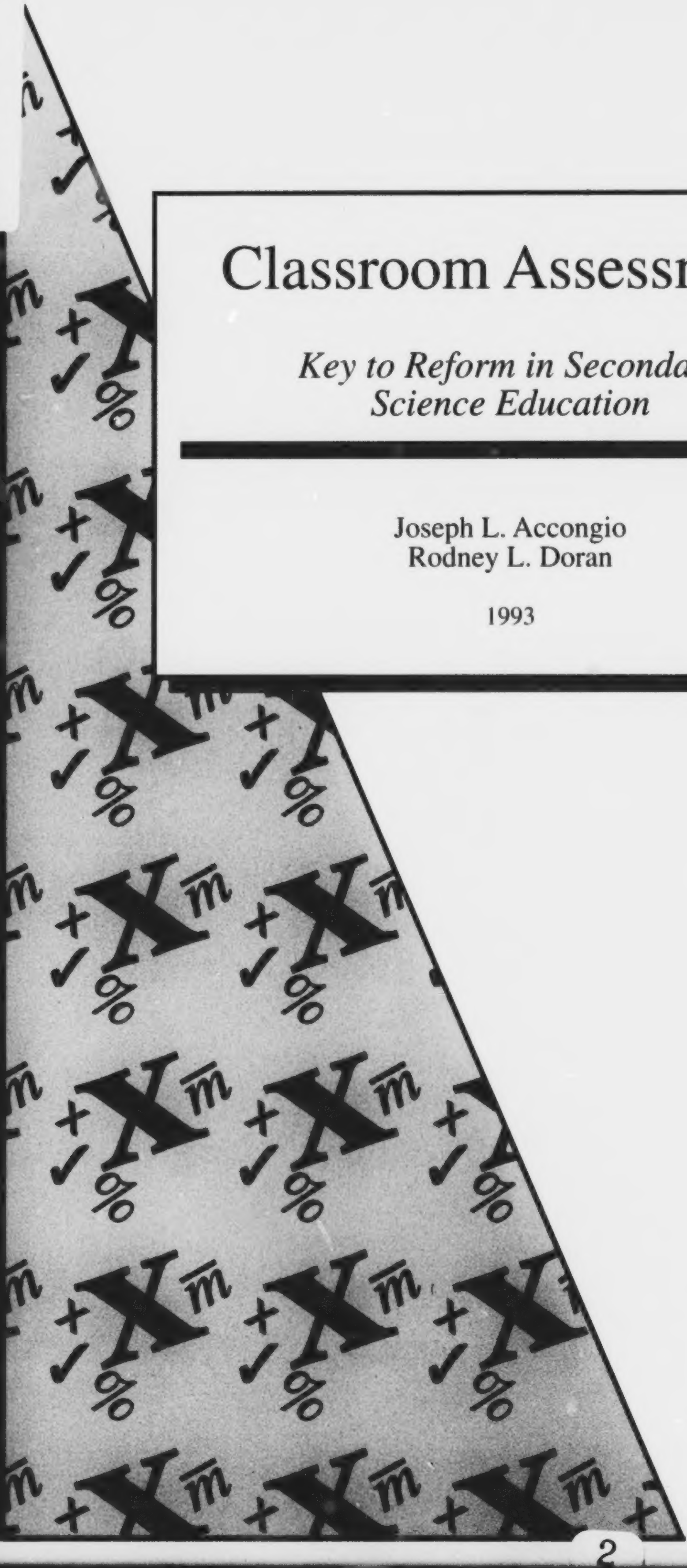
AUTHOR	Accongio, Joseph L.; Doran, Rodney L.
TITLE	Classroom Assessment: Key to Reform in Secondary Science Education.
PUB DATE	93
NOTE	208p.
AVAILABLE FROM	ERIC Clearinghouse for Science, Mathematics, and Environmental Education, 1929 Kenny Road, Columbus, OH 43210-1080.
PUB TYPE	Guides - Non-Classroom Use (055) -- Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC09 Plus Postage.
DESCRIPTORS	Educational Change; *Science Education; Science Process Skills; *Scientific Literacy; Secondary Education; *Student Evaluation
IDENTIFIERS	*Authentic Assessment; *Science Achievement

ABSTRACT

Science educators and other are faced with determining what changes must be made in education to insure scientific literacy for all students. One area that has been identified as playing a crucial role in science achievement outcome is assessment. This monograph focuses on improving assessment techniques used by teachers as they devise tests to measure and evaluate the outcomes of science instruction in their classrooms. The following chapters are included to provide insight on the role of assessment in science education reform: (1) "Assessing Science Achievement in the Middle Schools"; (2) "A Framework for Teaching and Assessing Science: The Nature of Science and the Nature of the Learner"; (3) "Assessing Levels of Cognition in the Content Areas of Science"; (4) "Assessing Process Skills: Scientific Thinking, Inquiry and Problem Solving"; and (5) "Authentic Assessment in Science: Performing Like a Scientist." The latter half of the document is composed of appendices, that contain illustrative items in the cognitive domain, scientific problem solving, and authentic assessment tasks proficiency profiles for science process skills. (ZWH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 370 774



Classroom Assessment

Key to Reform in Secondary Science Education

Joseph L. Accongio
Rodney L. Doran

1993

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Rodney L. Doran

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

BEST COPY AVAILABLE

Classroom Assessment

Key to Reform in Secondary Science Education

Joseph L. Accongio
Rodney L. Doran

Available from:
ERIC Clearinghouse for Science, Mathematics,
and Environmental Education
1929 Kenny Road
Columbus, Ohio 43210-1080
614-292-6717 (voice)
614-292-0263 (fax)
ericse@osu.edu (e-mail)

Table of Contents

Acknowledgments v

Preface 1

Chapter 1 Assessing Science Achievement in the Middle and High Schools

Purposes of Assessment 4
Challenges of Assessment 5
Improving Science Assessment 7

Chapter 2 A Framework for Teaching and Assessing Science: The Nature of Science and the Nature of the Learner

What is Science? 10
Purposes of Science Education 10
Developmental Levels of Cognition Among Learners 12
Learning Styles 15

Chapter 3 Assessing Levels of Cognition in the Content Areas of Science

Levels of Thinking 19
Applying Bloom's Taxonomy 20

Chapter 4 Assessing Process Skills: Scientifgic Thinking, Inquiry and Problem Solving

Critical Thinking in Science 23
A Framework for Assessing Critical Thinking in Science 25
Elements of Scientific Inquiry 26

Chapter 5 Authentic Assessment in Science: Performing Like a Scientist

What is Wrong with Current Assessment Practices 27
Building Performance Tasks in Science 29
Scoring Performance Tasks in Science 30
Measuring Proficiency in Science Process Skills 33
Psychomotor/Laboratory Assessment 34

Epilogue 39

References 40

Appendix A Illustrative Items in the Cognitive Domain 42

Appendix B Illustrative Items for the Scientific Problem Solving 69

Appendix C Proficiency Profiles for Science Process Skills 140

Appendix D Illustrative Authentic Assessment Tasks 152

Acknowledgments

The authors wish to acknowledge the support of Delta Inc. in the production of this monograph. We would like to thank Tom Richardson, David Butts, and James Shymansky for facilitating this support and Robert Howe for initial review of the manuscript for publication. Special thanks are extended to Stanley Helgeson for his encouragement and detailed editing. It is hoped that this publication will aid teachers and supervisors who are interested in expanding and improving their assessment efforts.

The authors would like to extend a special thanks to Barbara Alborano, Clara Moreno, and Patricia Kaiser for keyboarding and proofreading this manuscript.

To our wives and families, we would like to express our appreciation for encouraging us to complete this manuscript and for tolerating the many hours devoted to this task.

CLASSROOM ASSESSMENT: Key to Reform in Secondary Science Education

PREFACE

Few people read the preface, so we want to thank you for spending a few minutes on these pages. ***Classroom Assessment: Key to Reform in Secondary Science Education*** is about the teaching and testing of science in American schools today. The American Association for the Advancements of Science, in their recent publication, SCIENCE FOR ALL AMERICANS (1989), notes that scientific literacy has emerged as a major goal of education. However, general scientific literacy eludes us in the United States.

> "A cascade of recent studies has made it abundantly clear that by both national standards and world norms, U.S. education is failing to adequately educate too many students--and hence failing the nation. By all accounts, America has no more urgent priority than the reform of education in science, mathematics, and technology." (AAAS, 1989)

In spite of the overwhelming attention in the literature to the crisis in science education, there seems to be little progress in our classrooms. There still exists a gulf between the stated goals of science education and the measured outcomes of instruction. There are many reasons for this, to be sure, but perhaps inappropriate techniques of assessment--a crucial element of the curriculum-instruction-evaluation cycle--is one of the most compelling.

There has been a great deal written recently about assessment, measurement, testing, and evaluation. These words are used in many contexts and in a variety of ways by different people. This may lead to confusion and misinterpretation as one reads this monograph. The following paragraph from Doran, Lawrenz, and Hegelson (1993) may be helpful. They defined assessment as:

> "the collection of information, both quantitative and qualitative, that has been obtained through various tests, observations, and many other techniques (e.g., checklists, inventories, etc.) that are used to determine individual, group, or program performance. Measurement is thought to be a closely related term to assessment. However, measurement is viewed as not being quite as encompassing a term as assessment. Measurement has generally been defined as the process of testing, but it is also accepted as a more encompassing term that has included the use of observations, discussions, etc., as well as paper-and-pencil tests. The term testing is used to describe various teacher-made tests, as well as standardized forms of testing, such as inventories, questionnaires, and checklists. Evaluation is thought to be the process of making carefully determined value judgments and decisions related to the issues and concerns a given assessment as focused upon (e.g., students' achievement, program quality, etc.)."

This monograph focuses on improving assessment techniques used by teachers as they devise tests to measure and evaluate the outcomes of science instruction in their classrooms.

Student evaluation, in the form of grades we assign to students, is our way of "keeping score" in education. Class rank, a host of honors, college selection, and vocational opportunities depend directly upon the students' grade point averages. These grades, based almost exclusively on paper-and-pencil assessments and examinations, determine whether or not a student passes a course, whether or not he or she graduates, and ultimately impact on the quality of one's entire future.

Traditionally, grading in science courses focuses on how well students recall bits of information. While it has become commonplace to attack paper-and-pencil testing, it still represents a more than adequate technique for determining whether or not a student has mastered certain kinds of scientific knowledge, especially facts and terminology. However, written tests are not very effective in assessing whether or not a student can apply scientific knowledge in a real world context. Present assessment practices rarely include opportunities for students to perform activities related to the work of scientists and to be rated on those abilities.

Neither the massive curriculum projects of the 60s and 70s nor the numerous educational doomsday reports published over the past decade have significantly impacted on the overuse of teacher "monologue" methods in the science classroom. It has been estimated that more than 75% of secondary classroom teaching time is occupied by didactic teaching methods characterized by teachers talking. (Adler, 1982) This instructional strategy stresses the acquisition of information at the expense of the broader comprehensive objectives of science education. Unfortunately, the simple accumulation of discrete chunks of knowledge is the least durable form of learning. Most students begin to forget the content taught in such courses as soon as they have finished the final examination. Could you, at present, pass the examinations that enabled you to graduate from college even a few years ago? Except for the few who regularly use that knowledge, the answer most assuredly is no! Little wonder then that international comparisons place American students well below par on the scientific literacy scoreboard.

Why does an evaluation system, based on scoring the accumulation of facts, persist? What purpose does it serve? First of all, we are all familiar with the paper-and-pencil mode of assessment. This is a very efficient and arguably appropriate, way to measure students' ability to recall facts. However, this mode has become the favored format of assessment for all levels of learning--whether it is appropriate or not. It may be that classroom assessments consisting predominantly of recall items are perceived to be easy to construct, easy to score, and lead easily to psychometric analyses. Secondly, textbook companies and testing services provide an unending source of these items to model or to be used directly by teachers. Thirdly, this overuse of the paper-and-pencil format may be a direct result of the need to reduce complex learning and teaching behaviors to simple traits that can be converted to a number for the sake of assigning grades. Has the over-dependence on simplistic, convenient assessment procedures inhibited our teaching of science? More importantly, has our lack of attention to a more comprehensive assessment of the goals of science education restricted our students' potential to learn science?

It is profoundly important for science teachers to take as much care with test development and administration, and interpretation of collected data as they do in the preparation and delivery of their classroom instruction. Teachers must avoid the tendency to regard assessment as something to be done as quickly as possible after teaching has ended in order to "shake out a grade." The results of such practices inevitably lead to a misleading portrayal of the success of teaching in our science classrooms. High school transcripts tell us little about what permanent knowledge of science and science skills our students possess and even less about what they can actually do. Is it possible for a student to pass all of his or her science courses and still remain a scientifically illiterate American?

What must be done to correct this apparent contradiction between the stated goals and assessed outcomes of science education? As the learning behaviors become more complex and integrated, they are more difficult to describe, teach, and assess. The following recommendations are among the many needed changes in the way we function as science educators.

- We must agree on the relative emphasis to be placed on the various science education expectations in the schools and design assessment instruments to measure each appropriately.
- We must avoid the overuse and misapplication of expedient testing techniques.
- Teachers must be adequately trained in methods that effectively teach and evaluate desired behaviors at levels above recall and simple comprehension.
- Students should be expected to perform actual investigations dealing with scientific concepts, problems, puzzles, and phenomena. Evaluation should be based on their demonstrated abilities to do so--not simply on their knowledge of the process involved.

Acquiescence to expedient assessment practices is especially likely when the essential goals and priorities in science education are unclear or have been lost. The goals of developing higher levels of thinking, learning ways of generating knowledge in science, understanding complex concepts and perplexing phenomena, and being equipped to problem solve in science are vitally important components of scientific literacy for the 90s and beyond. Yet, we hardly measure these outcomes at all in the science class. We can no longer accept inappropriate and superficial assessment practices. It is axiomatic in education that the test steers instruction. There is nothing wrong with this reality if we have tests worth teaching toward. If we are to make science more meaningful for our students, a great deal depends on how well we mesh the focus and balance of classroom teaching, testing, and grading with the goals of science education.

This monograph was undertaken with the notion that the improvement of assessment practices directly relates to the improvement of instruction and achievement in science. If assessment is seen as integral to the educational enterprise, then the reform of science education can succeed, and the promise of scientific literacy for all America's students can be fulfilled.

CHAPTER 1: ASSESSING SCIENCE ACHIEVEMENT IN THE MIDDLE AND HIGH SCHOOLS

Introduction:

The more skeptical observers of today's schools suggest--often seriously--that American students spend more time being tested than being taught. Daniel and Lauren Resnick (1985) claim that our students may be the most tested but the least examined youngsters in the world. This may be an exaggeration, but no one can deny the enormous amount of testing in U.S. schools. Why is it then, that American students fare so poorly when compared to others around the globe?

We cannot minimize the influence that testing has on the lives of our students. Our methods of testing clearly affect the way we are teaching, and directly relate to what our children are learning. This in itself demands that the design, preparation, administration, and use of testing be raised to a level commensurate with the goals of science instruction. In this regard our teaching and testing practices are failing our children. If the reform of science education is to succeed reform in science assessment must happen now.

Purposes of Assessment

The assessment of science achievement plays a powerful role in American schools today. Data from tests are used to group and track students by ability, resulting in a wide diversity of learning experiences and opportunities. Teacher-made tests are used to diagnose knowledge and skill levels prior to instruction. Locally-developed and standardized tests are used to measure the attainment of instructional objectives and are used to evaluate teaching and instructional strategies.

Tests are intended to provide information about students' knowledge and abilities to that appropriate planning can take place. There are many ways to gain such information including observations, interviews, formal surveys, and other classroom-based techniques. Yet, the multiple-choice form of assessment has emerged as the most popularly accepted, and supposedly most objective, way to obtain all of this information. The teacher's responsible judgments (exercised or defaulted) are the most important influences on the quality of assessment, the interpretation of test results, and as a consequence, the students' future lives. To make these crucial judgments, teachers need a thorough knowledge of assessment and test construction.

Teaching science requires constant appraisal if for no other reason than to determine the success of that teaching. But, in the complex world of education, measurement of success must be carried out on many levels and from many perspectives to be most useful. While it goes without saying that every teacher needs to know what effect his/her instruction is having on students, several other reasons for assessing what goes on in the classroom are equally compelling.

Assessment provides insight to a student's rate of progress, reinforces effective learning habits, helps establish future time and effort commitments, and serves as proof of learning. The results of assessment can build student confidence or it can build anxiety. It guides, clarifies, instructs, motivates, or disenchants. For many students, getting a good grade on a test seems to be more important than the learning outcomes that the assessment is designed to measure.

Students need constant feedback so they can discover their own strengths and weaknesses. Through a well-articulated program of measurement and evaluation, students become aware of their natural aptitudes and begin to form the all-important subject-related self-concept. The student's image of him/herself (as a student of science) has a profound effect on future success in science, and whether or not the student chooses a career related to science. Feedback provides students with a view of themselves in regard to their standing and acceptance among their peers and satisfies personal and parental expectations.

Parents place a great deal of faith in teachers and in their judgments. Traditionally, the most parents receive in the way of reported evaluation is a letter grade, a percent, and one or two generalized comments--hardly a wealth of information about their children. Needless to say, evaluation reports must be accurate so parents can determine their children's progress. Not all parents can assess the learning abilities and deficiencies of their children. Valid and reliable knowledge about the capabilities of their children is vitally important to making informed choices regarding a child's future.

All children have anxieties, conflicts, problems, and needs that surface in the classroom. Parents often find it difficult to determine their role in assisting their children through these school-related situations. Parents depend almost exclusively on report card grades to help determine their involvement in their children's education, as well as to recognize successes and failures, and to take appropriate action. In this age of accountability, parents also look to their children's grade reports and standardized test results to help evaluate the teacher's competence and the strength of the local educational program.

Of course, elected school officials and their administrative staff members are extremely interested in the evaluation emanating from the classroom. Test data are often used to determine the strengths and weaknesses of teaching strategies and the instructional program. Beyond this, grades are used to identify curricular areas in need of concentration, correction, and revision. These results have implications on inservice training and other staff development initiatives. Summative report card grades have always been used as the main source for identifying student placement needs, leading to remediation, acceleration, enrichment, or alternative educational interventions.

School officials are also vitally concerned with the educational product presented to the tax-paying public. Test scores and grades are a concrete source of documentation comparing achievement and relative effectiveness among area schools. The success of the local educational program to some extent defines the image and quality of life in the community. The importance of appropriate evaluation extends to employers and agents of higher educational institutions who look to the recorded evaluative information as the major student profile component when selecting candidates.

Challenges of Assessment

Comprehensive assessment and well founded evaluation through paper-and-pencil techniques alone is very difficult, if not impossible to accomplish. To be valid and meaningful, assessment strategies must satisfy a multitude of conditions. Below are a few of the difficulties facing the educator in achieving a high-quality student assessment:

- Gathering unbiased evidence of student achievement.
- Designing assessments that are fair to all the students.
- Obtaining valid data.
- Finding or creating reliable assessments.
- Establishing effective methods of assessment that enhance learning.
- Designing and administering tests that are not punitive.
- Designing assessments that measure the degree of learning directly related to specific instructional objectives.
- Matching assessment formats to instructional purposes.
- Designing assessment items that measure learning in the various domains.
- Constructing assessments so they include many items that are above simple recall.
- Employing all levels of evaluation (diagnostic, formative, and summative) in the teaching cycle.
- Improving the interpretation and use of assessment data to enhance teaching and communication with parents.
- Designing "authentic" assessments that ask students to perform tasks related to what scientists actually do.

Tests--even the best of them--can never provide an absolute measure of a student's abilities, knowledge, or skills. Assessments are, by their nature, temporal estimates. Several factors beyond the difficulty of the knowledge being measured can prevent a student from doing well on a specific test on any given day. Here are some test-related frustrations facing students:

- The language of the directions, and the items themselves, may be vague, ambiguous, unfamiliar, or above the cognitive level of the student.
- The student may not understand the nature of the tasks to be performed.
- The student, because of the anxiety of the testing situation, may be unwilling to risk being wrong, and may "freeze."
- The student may be distracted from the test itself by the mechanics of marking an answer sheet designed for the teacher's convenience in scoring the test.

- The student's concentration and seriousness toward test-taking may be thrown off by the injection of humorous or nonsense choices.

- The student may not understand the "weight" of the test or parts within the test and may fail to devote the right amount of time and effort to each part.

- The student may be disoriented by extraneous items on the test that have not been taught or are there just to "fill time."

- The environmental conditions of the test site may be uncomfortable, distracting, or simply different from the conditions of the learning site.

- The student may not be feeling well on testing day. (Based on Delawter and Sherman, 1979)

Tests considered to be significant to science education goals should be designed to assess specific outcomes related to knowledge and abilities. To the extent that these are not assessed, the results are not useful. Results that seem to defy common sense, observation, and experience must be called into question. Students have the right to a just and fair evaluation.

Improving Science Assessment

A number of significant issues can be identified to improve the procedures for measuring the teaching and learning of science. These should cause the re-examination of the fundamental purposes and procedures used by educators to create their assessment instruments. They imply an emphasis on improving teacher-made tests so students are measured fairly and in wide-ranging areas of cognition. Criteria for building an effective science assessment system include:

- Authenticity: Developing assessments which focus on more realistic appraisal, of the attainment of the stated objectives of science teaching derived from real world counterparts, especially in the area of performance.

- Scoring Rubrics and Proficiency Profiles: Preparing carefully thought out sets of scoring criteria and indicators of performance that are known to the student and scaled for the purposes of measuring growth.

- Diagnosis and Prescription: Using evaluation more and more for diagnostic and prescriptive purposes, as well as grading, placement, and tracking of students.

- Formative Evaluation: Shifting emphasis toward the use of formative evaluation as it relates to the improvement of instruction and the evolution of the science curriculum.

- Higher Order Thinking: Paying more attention to the development of assessment formats that measure critical thinking and problem-solving skills.

- Habits of Mind: Emphasizing the need to include the values, attitudes, and skills directly related to a person's outlook on knowledge and learning as important educational outcomes that should be measured.
- Learning to Learn: Assessing the student's ability to learn within the context of science. That is, has the student internalized knowledge-generating processes and problem-solving strategies in science?
- World Class Standards: The reform of science education must constantly compare American science achievement with that of countries with acknowledged educational and economic strength.
- Methodology: Methods and techniques for preparing and administering science assessments must be in keeping with the best and most current principles of the psychology of learning and motivation, and knowledge of human development and social context.
- Test Bias: Gender, racial, cultural, and economic biases must be consciously excised from tests and testing environments.
- Test Anxiety: Efforts must be made to reduce "test anxiety."
- Criterion-Referenced Assessment: Assessment practices must shift from norm-referenced to criterion-referenced techniques to promote achievement for all students.

Conclusion

One of the reasons why we are in a state of crisis in science education is that science teachers give little attention to the measurement evaluation of the outcomes of instruction. This may not reflect a lack of skill or the desire to improve as much as it indicates the complexity of the problem. There are difficulties involved in assessing learning in science that make the task formidable for the classroom teacher, and for those who are responsible for the development and evaluation of educational programs.

One difficulty in assessing achievement in science is that school administrators, teachers, parents, students, the business community, and the public in general have failed to agree on the relative emphasis to be placed on the outcomes of science education. Given such confusion, it is understandable that the development of effective classroom assessment has lagged, and the use of ordinary paper-and-pencil items to assess only the most common outcomes of science instruction persists.

In defense of the science teacher, there is a serious paucity of high-quality performance-based instruments available for classroom use. The preparation and scoring of performance tasks is a time-consuming process. Alternative assessment instruments and techniques such as checklists, observations, one-on-one conferences, anecdotal record keeping, product evaluation, and portfolio assessment, to mention a few, are even more time-consuming--and are often denigrated as being too subjective, and therefore, unreliable.

Perhaps the greatest obstacle to better assessment is a result of the limitations of the classroom teacher's own knowledge of modern science, and the psychology of learning--and how these relate to assessment and evaluation. Science teachers who face a heavy and diversified teaching schedule, a proliferation of teaching objectives, new advisor-advisee roles, and other professional responsibilities are prevented from spending sufficient time to maintain and advance their knowledge of science, learning theory, and effective techniques of assessment.

A clearer vision is needed on the part of science educators regarding what American science students must achieve--and how one can determine that it has, in fact, been achieved. Collections of valid items, tasks, questions, and problems designed to measure the learning outcomes related to the vast and varied purposes of science education are needed. Such collections would assist teachers substantially in improving classroom assessments and provide valuable models for teachers attempting to construct their own assessments consistent with multi-dimensional instructional purposes. In addition, science teachers need a concise reference that will enhance the creation of authentic and effective classroom assessments. The remaining chapters of this monograph are dedicated to providing science teachers a framework for teaching and assessing significant content and processes in science along with concrete examples of items that satisfy these purposes.

CHAPTER 2: A FRAMEWORK FOR TEACHING AND ASSESSING SCIENCE: THE NATURE OF SCIENCE

What Is Science?

Science is perhaps the greatest intellectual and cultural achievement in the history of humankind. It is a creative process in which the human mind tries to solve nature's puzzles. The process begins with wondering, imagining, and hypothesizing. Ideas are challenged and thoughts are questioned in an open-ended attempt to make sense of the complexity of nature. As scientists pursue their work, they encounter controversy and inconsistency. But, like detectives hunting for clues, scientists relate one observation to another to build conceptual ideas. These ideas are connected to others until the interrelationships grow in structure and meaning to yield scientific theories and principles.

The questions that scientists ask deal not only with the puzzles of nature but with the problems of humans and society. The relevant questions are of the "what" and "how" variety because scientists can only describe, analyze, and predict. Science cannot answer the ultimate "why" questions which are matters for theology, metaphysics, and philosophy. A science teacher may validly ask, "What happens when paramecia encounter an iodine solution?", or "How do cilia help a paramecium move?", but not "Why do paramecia exist?" Finding answers to the what and how questions is the task of the pure scientists. What to do with these answers is the job of the applied scientists--the engineers and technologists.

The body of scientific knowledge can be thought of as an ever-expanding deeply piled carpet that changes over time. The fibers tied to the many junctions of the webbing are the specific facts of science (descriptions of the elements, the structure of the typical plant cell, and the instinctive behaviors of spiders, for instance). One cannot tell much about the pattern of the carpet from viewing a few individual fibers. Similarly, when viewed in isolation, specific facts do not reveal much about nature.

As more and more of these scientific "fibers" are brought into view, we begin to detect patterns and soon we can form generalizations--nature begins to make sense. Digging deeper, we discover the warp and the woof of the "carpet," the underlying structure of these models, principles, laws, and theories which are the awesome achievements of scientific thinking and investigation. It is through a process of active inquiry, pattern-seeking, recombination of ideas and problem solving that science evolves. These ways of learning about our world distinguish science from most other realms of meaning. (Phenix, 1964)

Purposes of Science Education

Defining science "for the public" has always been difficult. If we consider science to be not only what scientists know, but that which scientists do, then it becomes more understandable. Scientists believe that nature is comprehensible. They systematically confront it and try to map its secrets. This endeavor emanates from the interconnected constructs: the accumulated body of science knowledge organized in the form of facts, concepts, principles, laws, and theories; and the knowledge-generating processes, the discrete scientific skills which underlie scientific problem-solving strategies and critical/creative modes of thinking.

Historically, science education has drifted among a variety of curricular emphases. Differences between the traditional "content-centered" curriculum and the more recent "process-centered" orientation have particular importance when it comes to designing assessment items. The content of science resides in the form of facts, terms, definitions, formulas, concepts, theories, and principles. Such specific chunks of information, that can be memorized and recalled, are often the sources for items found on objective tests. However, detailed bits of information that are not understood in the context of scientific inquiry or connected to the fundamental ideas in science are tentative bits of inert knowledge at best. Just as a weaver needs a hook to tie his fibers to the carpet, so do children need "hooks" on which to hang new science knowledge. Failure to make these linkages explicit to learners usually means that new factual knowledge is extremely short-lived and the long-term purposes of science education are thwarted. Too often school science begins and ends with the teaching of many independent facts of science.

The processes of science, in contrast, are the intellectual skills that include observing, categorizing, measuring, inferring, predicting, hypothesizing, and experimenting. While content-based instruction tends to focus on giving students information to be learned, process-based instruction emphasizes student involvement in actively generating knowledge and solving problems. It is believed that, while students will forget most of the content information they assimilated, the critical process skills of science that are developed will persist and be transferable to other areas of inquiry as well. Of course, both emphases are important to the ultimate goals of science education. However, the traditional emphasis on content has left acquisition of the process skills relatively neglected.

Recalling that science can be defined as "what scientists do", it is clear that "science" must be thought of as an action verb. Science should be something that students do--more than simply read about or have a teacher talk about. Students must directly investigate nature at many levels and in many contexts. While getting teachers to acknowledge the necessity of providing such active learning opportunities has never been a problem, it has, however, been exceedingly difficult to make it consistently happen in the classroom. This has been blamed on a variety of excuses as explained in such comments as: "activity-based instruction is expensive," "it takes a great deal of planning and preparation," "it requires a strong laboratory background on the part of the teacher," "it presents classroom management and control problems," and excuses, "it is very inefficient to the purpose of covering the curriculum." These may represent perceived obstacles that get in the way of the regular employment of process-based instruction in the classroom. However, one realistic obstacle is the science teacher's lack of knowledge concerning appropriate and valid assessment and evaluation strategies for examining process skills.

Goethe said, "Knowing is not enough; we must apply. Willing is not enough; we must do." Herein lie the bases for building the framework for instruction in science. The purposes of science education should be to provide experiences whereby students have the opportunity to:

- acquire, and apply scientific knowledge, to better understand nature, and science-related puzzles, phenomena, and problems, and
- develop competency in performing the knowledge-generating process skills and problem solving strategies that characterize the scientific method of inquiry.

In this model, the task of the science teacher is to provide learning experiences that illuminate these goals. The science curriculum should encourage students to emulate the attitudes, attributes, and behaviors of scientists in performing open-ended investigations dealing with important, real-world scientific challenges.

Perhaps the most damaging criticisms of science teaching in America are that science teachers generally teach the way they were taught and that they do not consider the learning characteristics of their students. It has been shown that the predominant mode of teaching (lecture/demonstration) that most teachers experienced in their own education is no longer considered the preferred pedagogy for effectively teaching science today. Most secondary school students may not have the ability to consistently think in the abstract, and therefore, may be unable to derive formal generalizations about science on their own. Effective teachers draw on the growing body of knowledge about human learning to maximize their teaching efforts. Knowledge about the nature of learning and learners is no less important to successful science teaching than is a teacher's knowledge of science.

Developmental Levels of Cognition Among Learners

A commonly accepted notion is that the science teacher has not taught if the student has not learned. The difficulties related to teaching science are compounded by the differences in developmental levels of learning and behavior displayed among a given group of students. In order to teach science successfully, teachers must attend to a wide range of cognitive differences. These differences impact significantly on the students' ability to grasp concepts, the depth and breath of their comprehension, and the extent to which they can assimilate experiences and apply new knowledge.

Traditionally, teachers focus on quantitative differences among their students, primarily concerning themselves with the range of scores on tests. Teachers often ignore the qualitative differences that reveal the range of students' abilities to engage in higher-order thinking. This results in an inadequate assessment of the individual student's ability to learn. Differences in cognitive development represent a major obstacle in a teacher's quest for providing meaningful educational experiences. There are numerous studies that have examined the intellectual functioning of adolescents in secondary school years which bear upon the problems of teaching science. The familiar operant schema of intellectual functioning, theorized by Jean Piagét et al, represents the most comprehensive of such studies and provides a way to understand the processes involved in learning science.

For science educators, the crucial aspect of Piagét's stage theory of development involves the transition from what is called concrete to abstract or formal thinking among students. In the concrete stage, which is the prominent mode of thinking for young adolescent students (aged 11-15), understanding and reasoning are restricted to experience with actual objects. Those students who exhibit characteristics of concrete thought need to manipulate and interact with the objects and materials continually; they need direct verification of their findings. The transitional thinkers--who are moving from concrete to the formal levels--seem to have an enhanced understanding of what is being taught, but they need reinforcement through concrete experiences to verify their own evolving beliefs. In the formal stage, which is the expected mode of thinking for many--but certainly not all--older adolescents, reasoning becomes more and more unrestricted. They can handle the hypothetical, the imagined, and the possible. These students can predict consequences as they mentally play "what if" while they reason through a puzzle or problem.

It is important that science teachers contribute to their students' intellectual development by providing them with a sequence of educational experiences from concrete to abstract and appropriately matched instructional questions and classroom tests. Thereby providing students with opportunities to gain information and grasp scientific concepts and skills at a level appropriate to their stage of cognitive development. The transition from concrete to abstract thinking, and any confident application of it to science, is a slow and gradual process varying from student to student--and slower than we might expect. A British study (Shayer, 1976) indicated that between the ages of 12-14 years, only 10-20% of the population were capable of abstract thinking, although 50-80% were fairly advanced concrete thinkers. So one might expect less than one fifth of all incoming high school freshman to have the ability to function fully on an abstract level.

It is possible to detect the students' stage of thinking by observing the strategies they employ when coping with problems, noting how easily they appear to absorb new information, and by listening to the questions they ask and the answers they give. The following outline provides some comparative indicators that distinguish the characteristic behaviors between concrete and abstract thinkers in the context of science. These indicators may be useful in designing items that draw learners from where they are on the continuum of intellectual development to higher levels of thinking.

Indicators of the Stages of Cognitive Development in the Context of Science

I. Concepts

The concrete thinker

- is not able to grasp physical concepts with multiple variables such as volume and density.

The abstract thinker

- is able to understand physical concepts and can handle them mathematically.

II. Rules, Techniques, and Models

The concrete thinker

- is able to follow detailed procedural rules and understand concrete models.
- is not able to understand and employ theoretical models.

The abstract thinker

- is able to develop procedural rules and understand concrete and abstract applications.
- is able to understand and employ theoretical models.

III. Manipulating Variables

The concrete thinker

- has difficulty distinguishing between relevant and irrelevant information.
- has difficulty distinguishing between causes and effects.
- is not able to control variables.
- is limited in understanding two variable problems.

The abstract thinker

- can systematically handle relevant and irrelevant information.
- recognizes causes and effects in problems and experiments.
- will control variables properly.
- is able to handle problems involving two or more variables without confusion.

IV. Puzzle and Problem Solving

The concrete thinker

- can tackle puzzles and problems, provided that the actual objects and items can be handled and observed.
- is dependent on what is learned through the senses.
- bases judgments and conclusions solely on that which can be immediately verified.
- uses trial and error as the predominant method of investigation.
- is severely limited in logical and deductive reasoning.

The abstract thinker

- is free from the limitations of having to deal with physical objects and the here and now.
- can tackle theoretical problems.
- can propose hypotheses and evaluate them mentally.
- proposes solutions that are not only probable, but possible.
- is not restricted to the circumstantial when making judgments and drawing conclusions.
- employs logic, inductive, and deductive reasoning.
- seeks complete and detailed explanations.

Learning Styles

In addition to having to concern ourselves with a student's developmental level of cognition, it is equally important to consider the student's learning style. Some students learn best by hearing, others by seeing and yet others by touching and manipulating. The need of a student for a sequence of "hands-on" physical learning experiences does not necessarily indicate a learning deficit. On the contrary, it more often is an expression of a specific learning and processing style that must be accommodated to ensure success in the classroom.

The learning style (or learning modality preference) refers to a learner's preferred channel for receiving information through one or more of the senses. Major types of preferences include: visual, auditory, and kinesthetic/tactual. The following outline indicates several distinguishing characteristics among the various styles of learning. This information may be valuable in designing assessment items that measure student understanding unclouded by a mismatch between test format and a student's learning style.

Learning Style Indicators

I. Visual Learners

- choose books over an audio tape or film over discussion.
- may watch teachers' faces intently trying to lip read and use body language cues.
- may appear to be less intelligent than tests indicate.
- prefer to work alone rather than in an information-sharing situation.
- can recognize and match colors, objects, letters and words.
- are able to complete visual puzzles.
- recognize incomplete pictures, words, or letters.
- often respond in words or phrases rather than sentences.
- follow directions better when shown what to do rather than when being told.
- ask for repetition of oral statements.
- take cues from actions of other students during instruction.
- may have difficulty using simple verbs correctly.
- may have difficulty distinguishing and identifying sounds.
- may have difficulty isolating sound from impinging distracting sounds.
- may have difficulty rhyming words.
- may have speech problems.

Implications for designing effective instructional strategies and assessment items for visual learners:

- presenting information in a variety of ways (e.g., books, newspapers, films, TV, pictures, artwork, etc.)
- offering preferred seating arrangements which include sitting near a student with good auditory skills who can help the visual learner when confused about verbal directives.

II. Auditory Learners

- choose records or tapes over books and discussion over filmstrips.
- often seem brighter than tests show.
- score higher on tests when read aloud and asked to respond orally than when told to read items and write the answers.
- generally complete items on tests correctly but may do poorly due to lack of speed.
- recognize environmental sounds.
- follow verbal directions.
- understand stories read aloud.
- discriminate sounds and words.
- recognize incomplete, simple verbal analogies and rhymes.
- often skip sections of tests or books.
- read word by word.
- produce disorganized written work.
- frequently lose place when reading.
- fail to attend to written or visual tasks.
- seldom finish work.
- focus attention on sound when film or filmstrip is used. Interest in the visual presentation is secondary or nonexistent.

Implications for designing effective instructional strategies and assessment items for auditory learners:

- transmitting information in aural manner.
- providing seating location near the source of sound.

- amplifying sound.
- teaching reading phonically.
- utilizing lecture, panel and class discussions.
- using a marker to keep place when performing visual tasks.

III. Kinesthetic/Tactual Learners

- need concrete experiences rather than abstract learning.
- construct, draw, and act out.
- enjoy using and building models or projects for science but tend to avoid reading the text.
- are interested and stimulated by the laboratory aspects of science.
- generally do poorly on written tests but appear to know the content when doing projects.
- demonstrate good balance and coordination.
- easily manage physical education activities.
- usually have neat handwriting.
- do better than average art work.
- are good at imitating others (enjoy playing charades or other action games).
- give little attention to aural presentations.
- would rather perform a play than read or hear it.
- tend to do poorly on written tests.

Implications for designing effective instructional strategies and assessment items for kinesthetic/tactual learners:

- using a variety of hands-on materials (e.g., laboratory equipment, workshop tools, art supplies, manipulative materials, etc.)
- providing an opportunity for role playing.
- providing concrete experiences rather than abstract.
- providing opportunities to become actively involved in learning.

[Adapted from: Hasenstab, Blood, and Starke-Kobrin: 1982]

Summary

When teaching and assessing science, it is imperative to combine the knowledge of what makes an effective science program with an understanding of the implications of current research from the cognitive sciences. Reform in science education must be comprehensive--focusing on the learning needs of all students, as well as on the purposes of science education. It requires that science educators recognize the ways students receive, conceptualize, and process knowledge. We must analyze the requisites and components of science education and decide what are the most important goals to strive for. Then, we must attempt to effect a match among the way students learn, the way we teach, and the way we test the objectives of science. Only through such thoughtful efforts can we hope to transform our schools to ensure the attainment of scientific literacy for all.

CHAPTER 3: ASSESSING LEVELS OF COGNITION IN THE CONTENT AREAS OF SCIENCE

Levels of Thinking

There have been many attempts to describe the levels of human thinking. Several theories of learning, including those of Ausubel, Montessori, Bruner, Gagné, and Bloom, et al., have had a special impact on science education. By far, the most widely accepted theoretical scheme used by educators is that developed by Bloom and his colleagues.

Bloom, et al. (1956), classified educational outcomes into three major domains: cognitive, affective, and psychomotor; and subdivided the cognitive domain into the six well-known categories--knowledge, comprehension, application, analysis, synthesis, and evaluation. A brief discussion of each category of Bloom's Taxonomy follows.

Categories of Bloom's Taxonomy

KNOWLEDGE is the most elementary level of cognition. It values only the memorization of facts, definitions, conventions, principles, and theories. This tends to be the level most used by teachers in developing items for tests. Many teacher-made tests have been analyzed by experts who found that almost 80% of their test items were at this level (Fleming and Chambers, 1983). When a test item requires the students to give an answer that they can recall from memory, that item is classified as a knowledge-level question.

COMPREHENSION is the second major category of this hierarchy of thinking skills. At this level, students must show that they understand that knowledge committed to memory. Comprehension clearly goes beyond simple recall of information. Before students can comprehend or understand, they must first have learned certain facts and definitions. This implies that students must "know" before they can comprehend.

Three sub-categories of comprehension are translation, interpretation, and extrapolation. Translation is tested when a student is (a) required to paraphrase or explain a concept in his/her own words, (b) given a chemical formula and required to name the compound, and (c) given a diagram, sketch, or photo and required to name the parts and then tell their function. Interpretation is evident when students are able to not only explain a diagram or collection of data, but also identify similarities and differences that exist. Extrapolation is identified by the students' ability to go beyond the reported data. They can extend data in a table or on a graph and discuss results which would be obtained beyond those presented.

APPLICATION is the third major category in Bloom's taxonomy. Application of knowledge to new conditions assumes the student (a) has sufficient command of facts, terms, laws, concepts, processes, principles, and theories, and (b) understands or comprehends them. Students cannot apply what they do not know and understand. A difficulty arises at this point in using the taxonomy since what may be application for one student may be merely knowledge for another, depending on the nature of the student's experiences and the nature of the instruction provided. Test questions at the application level should incorporate problem situations that are novel and have not been covered in instruction or on assignments.

When engaged in ANALYSIS, students must be able to take a complex situation or set of information and separate it into its essential elements or constituent parts. Students must be able to recognize and discuss relationships between these elements. A student can demonstrate the ability to perform at the level of analysis if, when given a piece of research or description of an experiment, he/she can pick out the stated hypothesis, determine the assumptions upon which the research was based, and recognize the validity (or lack of validity) of the stated conclusions based on the data obtained.

SYNTHESIS is essentially the reverse of analysis. When a student is engaged in analysis, he/she studies the situation--part by part--and tries to discover relationships between the parts. When a student engages in synthesis, he/she is expected to put together a whole from many parts through the mental manipulation of facts, concepts, processes, and skills obtained from experiences. Synthesis is near the top of the taxonomy because it requires that students have mastered knowledge and can apply and analyze that knowledge.

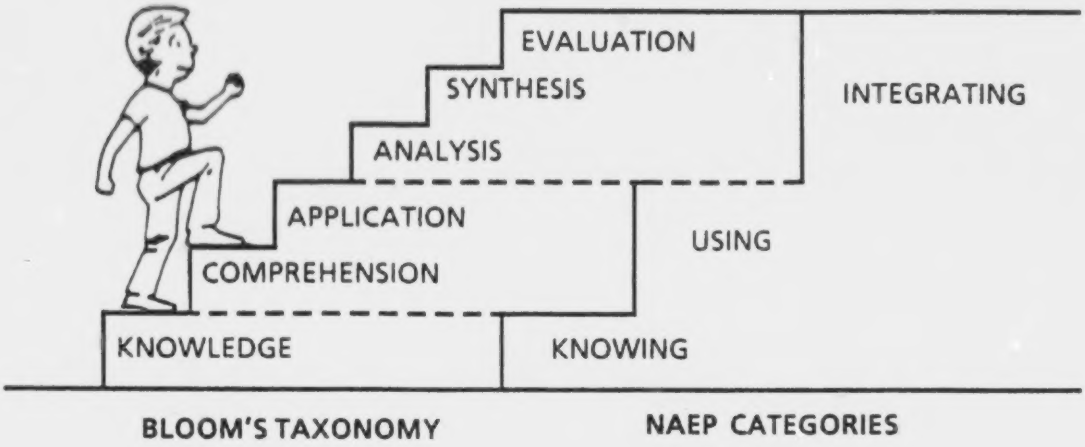
EVALUATION is the highest level in the taxonomy. Before legitimate judgments can be made on the value of some outcome of scientific endeavor, the evaluator must have a great deal of knowledge and understanding. Evaluation necessarily incorporates certain dimensions of the affective domain. However, evaluations must be distinguished from mere opinions which usually reside almost entirely in the affective domain. It should be clear that reasonable evaluations are based on all of the preceding levels of the cognitive domain.

Applying Bloom's Taxonomy

If a classification system is to be considered when creating instructional experiences and related tests, it must be in a form that is usable for teachers. It must have well delineated levels to allow the teacher to use it to ensure a variety of intellectual tasks, but it should not be so complex as to become burdensome to use. Many researchers and educators have found Bloom's six level taxonomy of the cognitive domain difficult to apply in practice. Any given test question may be answered in different ways by different students, depending on their specific past experiences with the topic, which tends to blur the divisions among the taxons.

Some educators have found it more satisfactory to use a simplified version of Bloom's taxonomy using three levels. A common approach is to retain knowledge (recall or recognition of specific information) as one category and to form two levels from the remaining categories. The second level usually includes Bloom's comprehension category and the capacity to perform routine, well-practiced applications of knowledge. The third level (encompassing analysis, synthesis, and evaluation) can be described as an integration of the higher order cognitive processes. The key feature of this highest level is the transfer of existing knowledge and skills to situations that the students have not previously encountered, to draw conclusions, and to determine the best approach to solving a problem. In either form, Bloom's cognitive scheme provides a useful hierarchical view of the various levels of cognition. It can be used effectively to devise instructional strategies, questioning techniques and assessment instruments.

In attempting to classify the dimensions of the cognitive domain in their "Framework for Science Assessment Exercises" (NAEP, 1986), the National Assessment for Education Progress devised what can be considered an abbreviated version of Bloom's taxonomy. NAEP defined three generic categories: "KNOWS," "USES," and "INTEGRATES." The following diagram illustrates this relationship.



Both taxonomies imply a hierarchy going from simple to complex categories of thinking. The simple behaviors are subsumed by the more complex ones.

In appendix A, there are sets of test items which assess science outcomes in the three levels of the cognitive domain proposed above. The specific purpose of each item is presented in a brief commentary following the item. The items selected are categorized by grade level, either middle school or high school, and by content area--earth science, biology, chemistry, and physics.

SET I. Assessement Items for the Cognitive Category--KNOWING

Successful performance on this set of items (#1-#8) depends on the student's ability to recall specific facts, concepts, principles and methods of science; to show familiarity with scientifc terminology; to identify these basic ideas in a variety of contexts; and to reorganize information into other words or another format. This category generally involves a one-step cognitive process.

SET II. Assessment Items for the Cognitive Category--USING

These items (#9-#16) test the student's ability to combine factual knowledge with rules, formulas, and algorithms for a specified purpose. Successful performance depends on the ability to apply basic scientific facts and principles to concrete situations; to interpret information of data using the basic ideas of the natural sciences; and to recognize relationships of concepts, facts, and principles to phenomena observed and data collected. This category generally involves a two-step cognitive process.

SET III. Assessment Items for the Cognitive Category--INTEGRATING

These items (#17-#24) test the student's ability to organize the component processes of problem solving and learning for the attainment of more complex goals. Successful performance depends on the ability to analyze a problem in a manner consistent with the body of scientific concepts and principles to organize a series of logical steps, to draw conclusions on the basis of available data, to evaluate the best procedure under specified conditions, and to employ other higher-order skills needed to reach the solution to a problem.

This category generally involves multi-step cognitive processes. In particular, it includes such mental processes as generalizing; hypothesizing; interpolating and extrapolating; reasoning by analogy, induction and deduction; and synthesizing and modeling.

CHAPTER 4: ASSESSING PROCESS SKILLS: THE ELEMENTS OF SCIENTIFIC THINKING, INQUIRY, AND PROBLEM SOLVING

John Dewey was among the first American educators to be concerned about the development of thinking skills in our schools' curriculum.

> "The conception underlying the school is that of a laboratory. It bears the same relation to the work in pedagogy that a laboratory bears to biology, physics, or chemistry. Like any such laboratory, it has two main purposes: (1) to exhibit, test, verify, and criticize theoretical statements and principles; (2) to add to the sum of facts and principles in its special line." (Dewey, 1896)

The need to plan and organize general education science experiences in such a way as to emphasize scientific thinking has long been part of American educational philosophy. One of the general education objectives of the 1947 presidential commission on higher education was stated as the need:

> ". . .to understand the common phenomena in one's physical environment, to apply habits of scientific thought to both personal and civic problems, and to appreciate the implications of scientific discoveries for human welfare." (President's Commission on Higher Education, 1947)

Few would argue that a foremost objective of education is to teach students to think and to solve problems. Yet from numerous national and international surveys, little evidence is found that critical thinking is being achieved. Rote learning exercises, physical activities requiring little or no thought, role-playing, lecturing, showing films, and videos seem to be the predominant modes of teaching. Little emphasis is being placed on developing students' ability to think. Why is there such an absence of teaching for thinking? This chapter will deal with the belief that the answer to this question involves the facts that 1) many teachers are unclear about the nature of critical thinking in science, 2) teachers lack experience in structuring classroom activities which develop thinking, and of equal importance, 3) they are unsure of how to assess the acquisition of such skills within the context of their subject areas.

Critical Thinking in Science

"Thinking" is the kind of word that has broad usage in the English language. Subtle differences among a vast array of applications of the word have caused confusion and led to the use of qualifying terms or alternative phrases. These include critical thinking, higher-order thinking, reflective thinking, inductive and deductive thinking, rational thinking, integrative thinking, logic, and creative thinking. Problem solving, decision-making, reasoning, and the scientific method are commonly used alternative phrases, each conveying different shades of meaning in a scientific context.

Dewey (1933) wrote that "reflective thinking" involves the following: 1) a state of doubt, hesitation, perplexity, mental difficulty, in which thinking originates, and, 2) an act of searching, hunting, and inquiring, to find materials that will resolve the doubt, settle and dispose of the perplexity. Watson and Glaser (1964) examined the concept of "critical thinking" and concluded that it was composed of attitudes, knowledge, and skills. Included are: 1) attitudes of inquiry that involve an ability to recognize the existence of problems and an acceptance of the general need for evidence in support of what is asserted to be true; 2) knowledge of the nature of valid inferences, abstractions, and generalizations in which the weight or accuracy of different kinds of evidence are logically determined; and 3) skills in employing and applying the above attitude and knowledge. After further research, Watson and Glaser identified five sub-abilities of critical thinking: 1) inferences; 2) recognition of assumptions; 3) deductions; 4) interpretations; and, 5) evaluation of arguments. These five abilities formed the sub-tests on their Watson-Glaser Critical Thinking Appraisal.

Dressel (1960) viewed critical thinking as a process that could be ordered in five steps: 1) recognizing and defining a problem; 2) clarifying the problem by making appropriate definitions, distinguishing between facts and assumptions, and collecting and organizing relevant information; 3) formulating possible explanations or solutions; 4) selecting one or more promising hypotheses for testing and verification; and 5) stating tentative conclusions. Dressel further noted that these steps are not necessarily considered discrete and sequential.

The steps in critical thinking most frequently cited are:

(1) Recognizing the problem
(2) Formulating an hypothesis
(3) Designing an experience or experiment to test the hypothesis
(4) Gathering pertinent facts or data
(5) Analyzing the facts or data
(6) Rejecting or accepting the hypothesis
(7) Drawing conclusions

One can see that these seven steps are the familiar sequence of steps in what is known as the scientific method. Critical (and reflective) thinking, and the scientific method are apparently closely related; many people see them as simply different names for the same intellectual activity (Skinner, 1983).

Critical thinking and scientific thinking involve abilities, attitudes, and skills. The methodology of science requires an attitude of inquiry, the utilization of process skills, and the application of knowledge of facts, principles, theories, abstractions, and generalizations from the content of science. Among the scientific attitudes and attributes are: being objective, anti-authoritarian, open-minded, and skeptical; being able to suspend judgment, and willing to adjust an opinion based on new findings; having the desire to understand the physical universe, and to seek experimental verification; and possessing curiosity, rationality, and intellectually honesty. Science proceeds through making careful observations, by using relevant knowledge to make hypotheses and inferences, and, finally, by seeking corroborative evidence.

A new emphasis on the scientific way of thinking is called for by the current reform in science education movement.

> "To ensure the scientific literacy of all students, curricula must be changed to reduce the sheer amount of material covered; to weaken or eliminate rigid subject-matter boundaries; to pay more attention to the connections among science, mathematics, and technology; to present the scientific endeavor as a social enterprise that strongly influences--and is influenced by--human thought and action; and to foster scientific ways of thinking." (AAAS, 1989)

The issue then is not whether critical thinking in science is important, but how these intellectual skills can be taught and assessed.

A Framework for Assessing Critical Thinking in Science

Clearly the view of science that is often neglected for assessment is that of science being an experimental subject concerned fundamentally with the intellectual processes required to effectively investigate and make sense of the natural world. Organizational schemes which reflect this view have been designed based on various conceptions of what is referred to as the scientific method. In all contexts, this endeavor addresses the planning of ways to generate hypotheses, make observations, and collect data. Organization, interpretation, and analysis of that data leads to an understanding of natural phenomena.

The Laboratory Structure and Task Analysis Inventory developed by Vincent Lunetta and Pinchas Tamir (1979) has been selected as an example of an organizational scheme that is characterized by its logical consistency in modeling the thinking that exemplifies scientific inquiry, and its well-defined categories for assessment. A brief discussion of each category of the inventory follows.

I. PLANNING AND DESIGN: The careful planning of investigations and experiments is fundamental to much scientific work. Experimental design depends on an appreciation of the scientific factors involved in a situation, as well as more general principles of experimentation, such as: the formulation of a question or definition of a problem to be investigated; prediction of experimental results; formulation of hypotheses to be tested; design of observations or measurement techniques; and design of experimental procedures, including the control of variables.

II. PERFORMANCE: Performing investigations and experiments is primarily concerned with the different procedures or skills used by students as they move from their perception of the problem to its solution. Many scientific processes can be viewed within a student's performance of a planned investigation. These include: making qualitative and quantitative observations or measurements; manipulation of, and technique in, using scientific apparatus; recording results of measurements and describing observations; performing numeric calculations; explaining or making a decision about experimental technique; and following through with the design of the investigation.

III. ANALYSIS AND INTERPRETATION: This category is concerned with the student's ability to examine and understand the collected data of an investigation or experiment. Behaviors that can be assessed in this category include the student's ability to: transform results into graphs and other forms for organizing and presenting data; determine qualitative and quantitative relationships; determine accuracy of experimental data; define or discuss limitations and/or assumptions that underlie an experiment; formulate or propose a generalization or model based on findings of the investigation or experiment; explain a relationship observed in the course of an investigation; and validate the hypothesis and formulate new questions or define a new problem based upon the results of an investigation or experiment.

IV. APPLICATION: This category focuses on the student's ability to apply scientific concepts and the findings of an investigation or experiment to novel situations. Given new information or a description of a new problem situation, students should be able to make predictions based on the results of previously experienced investigations or experiments; formulate reasonable hypotheses emanating from their new found knowledge; and apply appropriate experimental techniques that would lead to the solution of the problem.

Elements of Scientific Inquiry

For the purposes of this monograph, we have developed an abbreviated version of Lunetta and Tamir's Laboratory Structure and Task Analysis Inventory. The following is an outline of this scheme.

I. PLANNING: The student

- A. identifies a question, puzzle, or problem that can be investigated.
- B. makes predictions about an investigation.
- C. formulates testable hypotheses.
- D. designs an observation or measurement procedure.
- E. designs an investigation or experiment.

II. PERFORMING: The student

- A. carries out a qualitative observation.
- B. carries out a quantitative observation or measurement.
- C. manipulates apparatus.
- D. records observations.
- E. performs numeric calculations.
- F. critiques experimental procedures.

III. REASONING: The student

- A. organizes findings in tables and charts.
- B. graphs data.
- C. recognizes qualitative relationships.
- D. recognizes quantitative relationships.
- E. recognizes limitations and/or assumptions that affect the inquiry.
- F. draws conclusions or proposes models.
- G. explains relationships.

Appendix B is organized around the three major elements of scientific inquiry categories proposed above: Planning, Performing, and Reasoning. The items are further grouped by sub-category and grade level as indicated on the table below. Each item provided is followed by a brief discussion of its purposes, and a detailed scoring guide to clarify the assessment criteria for each. As the items are taken from various sources and serve several purposes, the number of points earned through each item will vary considerably.

CHAPTER 5: AUTHENTIC ASSESSMENT IN SCIENCE: PERFORMING LIKE A SCIENTIST

What is Wrong With Current Assessment Practices?

A major problem facing science educators today seems to be our inability to come to grips with what we would really like to have our students know, be able to do, and value. The problem is perpetuated by the way we report achievement to parents and to the public. High school report cards and transcripts tell us little about what students can actually do. Acceptance of standardized tests--driven by a desire to satisfy critics of public education, and the need to reduce complex intellectual behaviors to simple numbers--continue to define instruction. Within the prevailing framework for assessment, it is possible for a student to pass science courses and still remain scientifically illiterate. We must assess science ability in ways that take into account the fundamental purpose of science education--to enable students to use science to learn, to discover, to test, to solve, to predict, and to make sense of the universe.

Our current system of assessment continues to be dominated by paper and pencil testing techniques that reduce performance to the selection of one right answer. The problem with current teaching and testing practices is that they frequently reinforce the idea that getting the "right answer" is an adequate sign of ability. Rarely are students evaluated on their ability to understand, to attack, and to solve authentic problems in science from a performance stand point. "Authentic evaluation of educational achievement directly measures actual performance in the subject area. Standardized multiple-choice tests, on the other hand, measure test-taking skills directly, and everything else either indirectly or not at all." (Pett, 1990) What accounts for this kind of assessment taking hold so dominantly in America? Daniel Resnick, in his address to the Mitsushita Foundation Authentic Assessment Conference (1990), responded to this question as follows.

During the period of 1910-1930, associationists and behavioral psychologists such as Thorndike developed a theory of assessing knowledge claiming that the way people came to know things was by making associations between different kinds of perceptions and experience. Knowledge was considered to be the accumulation of these bonds. Thorndike viewed the world of knowledge as a machine in which there were all sorts of component parts. At the time, this was considered a valid mapping of the way in which knowledge was actually constructed. If you could show your mastery of part of the whole, you were demonstrating as much as could be shown by anybody, by that method of assessment. What we usually assess through the multiple choice test format is fragmented knowledge--fragments that do not cohere in any way.

Two underlying beliefs have contributed greatly to the wide use of the multiple choice and short answer assessment techniques, and have worked against the development of broader, more comprehensive performance tests in science. They are the "decomposability" and the "decontextualization" of knowledge and skills. Decomposability assumes that we can tease knowledge apart, separating it into a lot of different pieces without jeopardizing understanding and applicability. Comprehensive performance will take place at another place and time--perhaps in the work place. This view indicates that teaching and testing of separate skills is not only possible but valid. The idea of decomposability does not describe a paradigm that matches the work of cognition and psychology that has evolved recent decades (Resnick, 1990).

The current notion of how learning takes place is that even the most basic skills are acts of construction of knowledge. Based on this notion, it is vitally important for teachers to use assessments that demand students to demonstrate or perform all their competencies, knowledge, and higher-order thinking skills--acquired from a variety of learning experiences--to provide a true evaluation of their abilities. It is becoming increasingly clear that we must include comprehensive performance testing while students are still in public education--before they encounter the demands of the work place.

What are the implications of this change of paradigms? Efforts to assess thinking and problem solving abilities by identifying separate components of these abilities and testing them independently, will interfere with the actual teaching of such abilities. Assessing separate components has encouraged the overuse of exercises in which isolated components are practiced. Since the components do not enhance integrative thinking and problem-solving, students who only practice the components are unlikely to do problem-solving or interpretive thinking. The first element of this paradigm--decomposable knowledge--does not fit the current view of knowledge as a complex set of interactive skills tied to subject area information that we hold today (Resnick, 1990).

The second problem of the associationist and behaviorist psychologists' paradigm that doesn't work is decontextualization. The prevailing theory assumed that one could assess separate skills in any context. If a student could distinguish fact from opinion in one context, he could do so under all other contexts. However, we now know that we cannot teach a skill component in one setting and expect it to be automatically applied in another. Based on current cognitive psychology, we cannot validly assess an understanding in a context very different from the context in which it is practiced (Resnick, 1990).

Traditionally, American assessment attempts to make the complex simple by dividing learning into isolated and simple tasks for students to do; assuming that the students need not practice the "true test" of performance--the test of putting all the elements together. This seemingly logical approach of breaking tasks down into their components leads to tests that assess artificially isolated outcomes and provide no stimulation of comprehensive intellectual progress. Most fixed response and multiple choice tests reinforce this approach (Wiggins, 1989).

The belief in the underlying assumptions of decomposition and decontextualization permeates much of the science curricula and assessment practices used in our classrooms. However, many educators are trying to move away from those assumptions to more integrative and more comprehensive characterizations of what constitutes authentic teaching and learning. To be "authentic," achievement tests must be thought of as the intellectual equivalent of public performances. In science, the most likely authentic context for evaluating the teaching and learning that precedes such assessments is the "investigation". Many investigations are performed in the science laboratory. Others may be done in "regular" classrooms, parking lots, gymnasiums, parks, and other outdoor settings. What is unique is not the setting, but what the students are expected to do there--perform like scientists! Many of the specific skills and understandings illustrated in previous chapters of this monograph can be assessed during an investigation. Similarly, the elusive reasoning and integrating levels--with their broader, more demanding expectations--can be taught and assessed through investigations. The assessment of performance in the context of real-world, problem-based investigations can be described as "authentic assessment."

Building Performance Tasks in Science

As you begin developing performance tasks, the following suggestions adapted from Baron (1989) might be very helpful. Baron and her colleagues have developed a variety of performance tasks for use in science and mathematics. One task from this work will be illustrated later in this chapter. The development process begins with the underlying idea revised through drafts, trials, and includes the important scoring guidelines.

The Performance Assessment Task Development Process

Start with an Idea	● From a textbook or other book ● From a newspaper or magazine article ● From a life experience ● From conversation with colleagues or students
Test the Idea	● Does it center on an important concept, skill, or issue in science? ● Does it have meaningful context for students? ● Does it tie the concept or issue to real life? ● Does it lead students to deal with the concept or issue instead of just memorizing it? ● Does it make students use it, understand it, explain it to others, or otherwise take some ownership of it?
Begin Converting the Idea into a Performance Assessment Task	● Clarify the objectives of the task. ● What knowledge/skills/abilities/attitudes/attributes will students have to display in order to successfully handle this task? ● Write a complete guide statement, including task statement, purpose, and suggestions to students on strategy and focus. Keep your original objectives in mind throughout!
Consider Embellishments	● Can the problem be made "multi-faceted" (e.g., multiple performances or products around the same theme or experience)? Consider written exercises or reports, oral reports, group discussions, or performances, student logs or portfolios, self-assessments, etc. ● Can the task be structured to also elicit attitudes and attributes which can be measured (e.g., group cooperation, persistence, resourcefulness, etc.)? ● Can the task be structured to include group activity?

Consider What a Teacher Will Need to Know to Administer the Performance Assessment Task

- How does this task fit into the curriculum?
- What needs to be taught before the performance assessment test is administered?
- What materials and equipment are needed?
- What problems or difficulties are likely to occur?
- What kinds of assistance or intervention should the teacher be prepared to provide? What kinds of assistance should the teacher not provide? How should such interventions be treated in scoring?
- Develop NOTES TO TEACHER to include all of the above.

Design a Scoring Procedure for the Problem

- Consider your original objectives--how will they be revealed in student's responses?
- Decide whether you are assessing processes or products.
- Identify either dimensions of performance or aspects of the product which (a) reflect the objectives you had for the task, and (b) can be observed and rated with reasonable objectivity.
- Weigh the dimensions in proportion to their importance, using your own judgment and that of colleagues.
- Develop levels of performance which you feel are likely to be present in student performance or products.
- Determine the range of points to be allotted to each level of performance.
- Build a section within the guide to communicate to students how their performance will be evaluated.

Try the Performance Assessment Task Out

- Have one or more colleagues review it "cold" and critique it.
- Administer it in the classroom, in a low-stakes setting, to evaluate the task, not the students.
- Get feedback from your students:
  - What was good or bad about the task?
  - What would improve it?
- Try out the scoring system:
  - What was good/bad about it?
  - How can it be enriched by the examples of performance you have now collected?

Revise Accordingly Based on Your Experience

- Revise guide.
- Revise scoring system, and illustrate.
- Expand NOTES TO TEACHER.

Scoring Performance Tasks in Science

A critical need for assessments that will be accepted and used by teachers, parents, and students is a clear, concise, consistent scoring system. A scoring system is easier to use if it is based on a simple understandable schema of objectives or goals. At several places, we have used an organizational taxonomy of laboratory skills by Lunetta and colleagues. These researchers have also presented guidelines for rating student performances on the several categories within their taxonomy. We have included part of their scoring suggestions here. The detailed criteria for various levels of performance of the planning and design category are illustrated below. The scoring for the first category (Planning and Design) begins with "one-two," which corresponds to the most minimum skills, and ends with "nine-ten," which would represent the ideal/complete performance expected with students within some stage of sophistication/development. For the other four categories, only the "ideal" level is illustrated. One can develop the criteria for the other levels based on the Planning and Design example. Variation of this scoring model will be cited in illustrative tasks that follow.

A. Planning and Design

SCORE	CRITERIA
9-10	Able to present a comprehensive plan for investigation. Plan is clear, concise, and complete. Able to discuss plan for experiment critically.
7-8	Good, well-presented, detailed plan, but needs some minor modification. Understands overall approach to problem.
5-6	Plan is general with specifics barely adequate. Not a very critical approach to problem.
3-4	Relatively ineffective plan, needing considerable modification. Does not consider important constraints and variables.
1-2	Little idea of how to tackle the problem. Much help is needed.
N.O.	No opportunity to use these skills.

B. Manipulative Skills and Conduct of Experiment

SCORE	CRITERIA
9-10	Excellent performance of full range of skills. Appreciates precision of apparatus. Quantitative results within expected range. Can carry out plan in a reasonable time modifying it creatively effectively when appropriate.

C. Observations and Recording of Data

SCORE	CRITERIA
9-10	Correct observations specified, unexpected results noted. Errors and possible inaccuracies are not ignored. All relevant information accurately recorded in an appropriate form.

D. Interpretations of Data and the Experiment

SCORE	CRITERIA
9-10	Excellent handling of data, appreciation of error, and limitations of experiment. Knows when to search for additional information. Good analytical approach. Good appreciation of scale of experiment. Accurately calculates results from experimental data. Good written and oral appreciation of results and variables. Makes inferences from data to suggest explanations, relates the results of the experiment to other problems.

E. Responsibility, Initiative, Work Habits

SCORE	CRITERIA
9-10	Self-reliant, able to work with little supervision. Willing to tackle problems. Can work as part of a team, as well as on own. Safety conscious. Willing to help in running of laboratory if asked. Consistent and perseveres. Tackles practical with with enthusiasm.

(Adapted from: Lunetta, Vincent N., Avi Hofstein, Geoffrey Giddings, "Evaluating Science Laboratory Skills." The Science Teacher: NSTA, January, 1981.)

Measuring Proficiency In Science Process Skills

In work being done under the direction of Dr. Jean B. Slattery, Supervising Director of Curriculum Development and Support in the City School District of Rochester, New York, scoring rubrics for performance assessment tasks and proficiency profiles in science are being prepared for field testing. The following science process skill profiles, adapted by Accongio (1991), are based on the work of SAPA (Science, A Process Approach). These have been prepared as part of the work in Rochester's Goals, Outcomes, Measures, and Standards initiative for the K-12 educational program.

In order for students to better explore and understand nature for themselves, they must learn an array of inquiry skills used by scientists. These include the basic and higher level integrated process skills. The proficiency levels within the following set of profiles are ordered developmentally. That is, each level of proficiency subsumes the abilities listed in the levels preceding it. **Appendix C includes** profiles in the following areas:

A. Observing
B. Classifying
C. Measuring
D. Inferring
E. Predicting
F. Experimenting
G. Constructing and Interpreting Models
H. Questioning and Hypothesizing
I. Interpreting Data

This set of profiles is followed by a Proficiency Profile Rating sheet for use in the classroom with individual students.

Psychomotor/Laboratory Assessment

In the early 1980s, the Industrial Management Council of Rochester, New York, funded PRIS²M (Program for Rochester to Interest Students in Science and Math) in an effort to increase the achievement of minority students in math and science.

The overall goals of the PRIS²M program included the nurturing of specific laboratory skills. Familiarity with, manipulation of, and proper application of such equipment, tools, and apparatus as microscopes, calculators, bunsen burners, computers, graduated glassware, balances, and rulers require a certain level of manual dexterity, knowledge, and specific skills. The PRIS²M curricula sought to enhance students' science skills through a materials-intensive, hands-on approach to instruction.

This psychomotor performance assessment task, created by Accongio (1983) provides a measure of several crucial skills (measuring, graphing, tabulating, and constructing) that may be predictive of a student's future ability in laboratory situations.

INDUSTRIAL MANAGEMENT COUNCIL

ROCHESTER CITY SCHOOL DISTRICT

PRIS^{2}M JUNIOR HIGH SCHOOL CURRICULUM

PSYCHOMOTOR ASSESSMENT

Introduction:
The following activity is designed to see how well you work with materials and how well you follow directions. Remember, you cannot fail this test. Take your time, read carefully, and try your best to do what is asked of you. (40 minutes)

Materials Required: sharp pencil, metric ruler, scissors, adhesive label

Procedure: Work through the 12 STEPS of this activity.

STEP 1. Fill in the information required.

Name ______________________ I.D. Number ____________

Teacher ______________________ School ______________

Date ______________________ Class Period ____________

STEP 2. Fill in this same information on the heavier piece of paper with all the dots, numbers, and lines on it. (From now on, that paper will be called the CARD.)

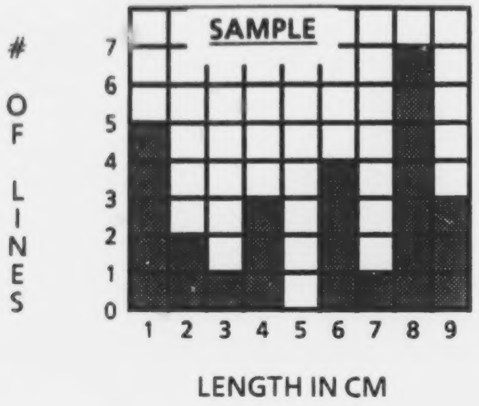
STEP 3. Connect dots on the CARD as indicated in the list below. Measure each line that you draw to the nearest centimeter.

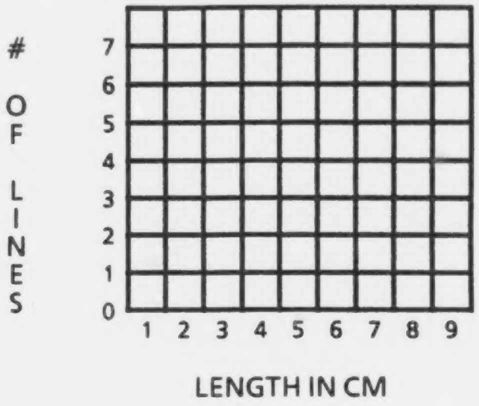
Connect dots 2 & 3. This line is about ______ cm long.
Connect dots 4 & 29. This line is about ______ cm long.
Connect dots 7 & 8. This line is about ______ cm long.
Connect dots 8 & 30. This line is about ______ cm long.
Connect dots 30 & 31. This line is about ______ cm long.
Connect dots 9 & 10. This line is about ______ cm long.
Connect dots 10 & 11. This line is about ______ cm long.
Connect dots 12 & 13. This line is about ______ cm long.
Connect dots 13 & 14. This line is about ______ cm long.
Connect dots 15 & 25. This line is about ______ cm long.
Connect dots 21 & 22. This line is about ______ cm long.
Connect dots 16 & 24. This line is about ______ cm long.
Connect dots 17 & 18. This line is about ______ cm long.
Connect dots 19 & 23. This line is about ______ cm long.
Connect dots 23 & 30. This line is about ______ cm long.
Connect dots 20 & 26. This line is about ______ cm long.
Connect dots 26 & 27. This line is about ______ cm long.
Connect dots 28 & 1. This line is about ______ cm long.

STEP 4. Count how many lines that you drew that were about 1 cm long and record that number on the chart found on the CARD.
Do the same thing for the rest of the line lengths given on the chart.

STEP 5. Graphing is one of the most useful ways of showing information. The sample graph below is called a BAR graph. The numbers listed along the bottom of the graph could be the lengths of the lines that you have just drawn. The numbers on the side represent the number of times that the same length of line was drawn.

Use the information that you recorded in STEP 4 to make your own graph on the blank grid below.





STEP 6. Cut along every solid line on the CARD except for the lines that you wrote information on, and the lines that make up the chart.

STEP 7. Fold along every dotted line. The dotted lines should remain on the outside of the fold.

STEP 8. Unfold the CARD and lay it flat on your desk. Be sure that no parts are tucked under. Notice that there are four arrows on the CARD. Two arrows at the left side marked "a" and "c" and two arrows at the right side marked "b" and "d".

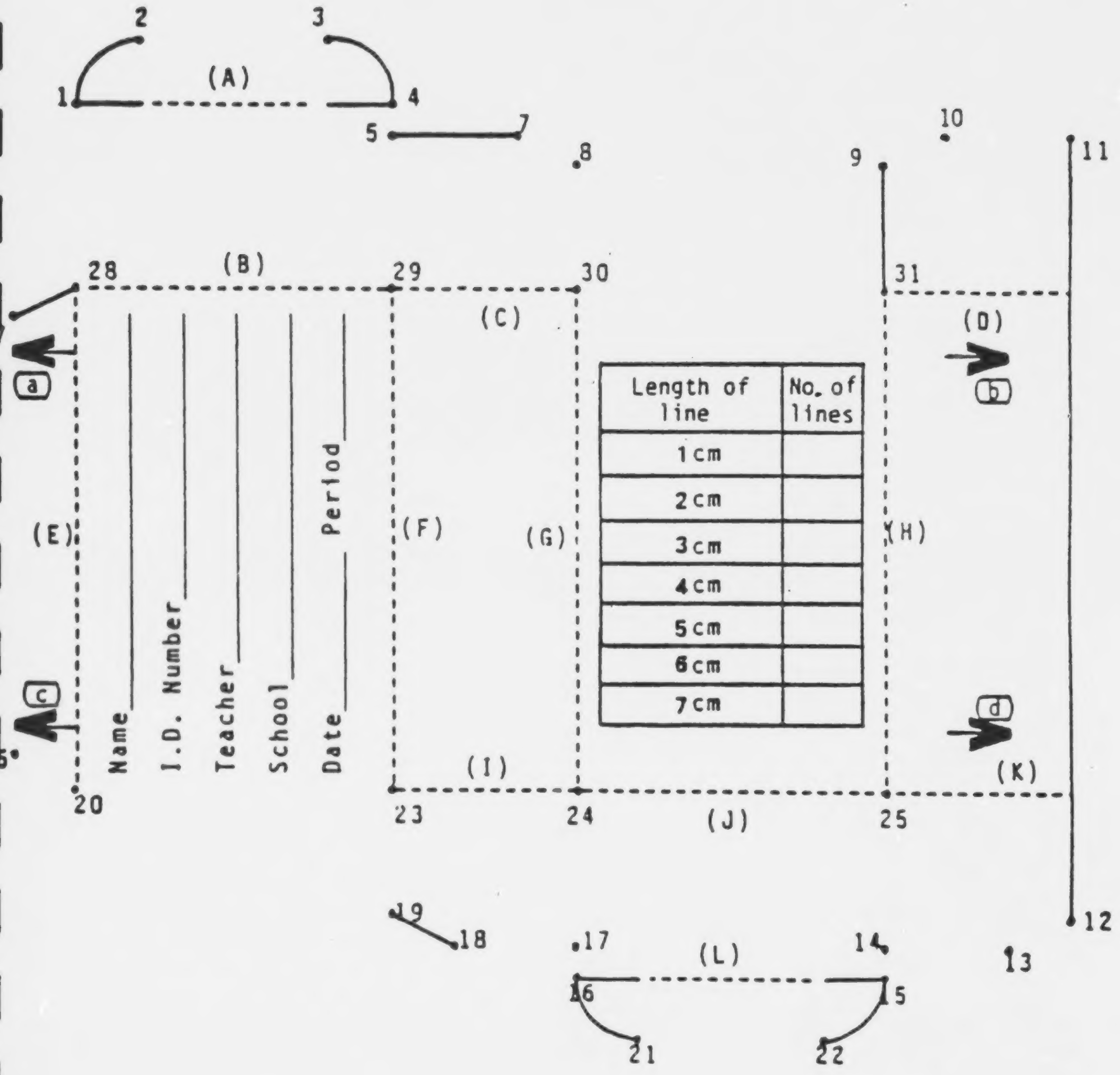
STEP 9. Fold the CARD along dotted line (G). Place the folded CARD on your desk so that you can see the section containing the chart and the arrows "b" and "d".

STEP 10. Fold over the flap showing arrows "a" and "c". Line up arrow "a" with arrow "b", and arrow "c" with arrow "d". Tape the flap in place so that the tips of the arrows are touching.

STEP 11. Pick up the CARD. Examine it closely. Handle it until you recognize what you have made. Finish putting this familiar object together.

STEP 12. When you are satisfied with your work, raise your hand and your teacher will collect your materials.

-- END OF TEST --


1
2
(A)
3
4
5
7
8
9
10
11
28
(B)
29
30
31
27
(C)
(D)
a
b
(E)
(F)
(G)
(H)
c
d
26
(I)
(K)
20
23
24
(J)
25
19
18
17
(L)
14
12
13
16
15
21
22

Name ____

I.D. Number ____

Teacher ____

School ____

Date ____ Period ____

Length of line	No. of lines
1 cm	
2 cm	
3 cm	
4 cm	
5 cm	
6 cm	
7 cm	

(NOTE TO TEACHER: COPY ON CARD STOCK)

Scoring Guide

The major portion of this assessment task examines students' ability to properly use a metric ruler to measure 18 lines accurately to the nearest centimeter. This subtest not only evaluates students' knowledge of metric measurement and estimation, but tests their ability to persist in a repetitive data gathering and recording process. Upon completion of the initial measurements, the students are asked to compile the gathered data and display it on a data table showing frequency of individual line measurements, thereby, assessing the lab skills of observing, counting, organizing, and tabulating data. The tabulation is followed by another subtest, the creation of a histogram. Finally, a set of detailed procedures are presented to the students that lead to the "construction of a familiar item" (a small box). Here, the students' ability to follow directions carefully and to manipulate materials are measured by their accuracy in folding, cutting, taping, and inferring the final step in the construction of a model box.

The point values for each subtest were determined by the number of components involved in the performance of each individual task. A factor was established for each subtest based on the following list of specific operations:

MEASURING

a) locate points
b) use straight edge
c) use ruler to measure line
d) record measurement

TABULATING

a) locate common line lengths
b) compile frequencies
c) record frequencies

GRAPHING

a) interpret data table
b) locate line intervals on horizontal axis
c) locate corresponding frequencies on vertical axis
d) plot graph (fill-in bars)

CONSTRUCTING

a) perform operation indicated using scissors and tape (cutting, folding, manipulating, taping)

RECORDING

a) recording pertinent information on the answer sheet and on the cardboard box

The common function on each subtest was "read and follow directions" and was therefore eliminated from consideration as a separate factor in scoring.

	Number of Sub-Factors		No. of Items		Weighted Subtest
MEASURING	4	X	18	=	72
TABULATING	3	X	7	=	21
GRAPHING	4	X	5	=	20
CONSTRUCTING	1	X	40	=	40
RECORDING	1	X	12	=	12
			Total possible score	=	165 pts.

Factors yielded total weighted scores when multiplied by the number of individual items composing each subtest. Total possible score for the entire psychomotor assessment is therefore, 165 points.

Illustrative Performance Tasks

The remainder of this chapter is organized around illustrative performance tasks grouped into three categories. The categories are determined by the mode of administration: "Guided Investigation," "Practical Investigation," and "Extended Investigation." Each of these assessment modes is described below.

In ***Guided Investigations***, one student works at a lab station with a test administrator who monitors and interacts with the student. This one-on-one procedure requires a considerable commitment of space and time from the students and adults. It must be carefully scheduled to obtain maximum impact from its use.

In ***Practical Investigations***, individuals or pairs of students perform with no intervention from a test administrator. Students work at lab stations where the equipment and material for their specific experiment are located. A number of different investigations from the same subject area can be prepared within one lab room. For instance, in high school chemistry, one could have separate set-ups for the following investigations: acid-base titration, reaction rates, mole concept, solubility, melting point, and hydrated salts. Investigations with the same experimental set-up are not located in adjoining lab stations.

The ***Extended Investigation*** occurs with individuals and/or groups of students at different phases or stages over several class periods. This method of assessment has been described as being "close to instruction" and more realistic in terms of how problems are commonly encountered and solved in the world beyond the classroom.

Examples are included for each mode of investigation. As these assessment tasks are complex in terms of student sheets, equipment lists, and detailed teacher notes and checklists, we will describe only some of these elements for some tasks. The reader is encouraged to contact the authors of this monograph or the respective developers directly for the complete set of information for trying some of these tasks. As always, the reader is encouraged to modify these illustrative items to fit other teaching/learning situations and to develop new tasks based on the items modeled in Appendix 12.

Epilogue

The last decade has witnessed a resurgence of interest in the assessment of science outcomes. A recent AAAS forum focused on "Assessment in the Service of Instruction." A major thrust of the National Center for Improving Science Education (the Network and BSCS) has been on assessment. Articles have proposed attention on performance assessment, portfolios, and higher order thinking processes. The results and reports of several national and international surveys have illustrated the importance of assessment.

Despite all the reports, forums, and surveys, the locus of assessment that has the most impact is that which is organized at the classroom and school levels. This assessment is accomplished without the large budgets and staffs engaged by those groups mentioned above. The authors realize the minimal support and moderate training that science teachers and supervisors have for their crucial assessment responsibilities. By no stretch of the imagination, does this monograph supplant the vast resources that would be needed to raise science assessment to the desired level. However, we believe strongly in the creativity, energy, and resourcefulness of the science teachers. We hope that by organizing and illustrating some top-notch examples of assessment items for a few key outcomes of middle and high school science programs that teachers will be able to extend and expand their repertoire of testing skills. It has been to this end that we have produced this volume.

Selected References

Accongio, Joseph L. Proficiency Profiles for Science Process Skills. Goals, Outcomes, Measures, and Standards Initiatives. City School District of Rochester, N.Y. Unpublished document: September 1991.

Accongio, Joseph L., and Rodney Doran. PRIS²M Evaluation Report, Industrial Management Council. Rochester, NY, 1983.

Adler, Mortimer J. The Paideia Proposal. New York: Macmillan Publishing Co., 1982.

American Association for the Advancement of Science. Science for All Americans. Washington, D.C.: American Association for the Advancement of Science, Inc., 1989.

Baron, et al. Common Core of Learning Assessment Project. Connecticut State Department of Education. Division of Research, Evaluation, and Assessment, July 31, 1989.

Bloom, B.S. (editor). A Taxonomy of Educational Objectives: Handbook I, The Cognitive Domain. New York: Longman, 1956.

Delawter, Jayne, and Barry W. Sherman. "Skirting the Perils of Testing." Learning, October, 1979, Vol. 2, pp. 99-102.

Dewey, John. How We Think. Boston: D.C. Heath and Company, 1933.

Dewey, John. "University Record." 1, 32, Nov. 6, 1896, pp. 417-422.

Doran, Rodney, and Willard J. Jacobson. Science Achievement in the U.S. and Sixteen Other Countries. New York: Columbia University, 1988.

Doran, Rodney L., Frances Lawrenz, and Stanley Helgeson. Assessment in Science. Unpublished manuscript: 1993.

Dressel, P.L. "How the Individual Learns Science." Rethinking Science Education. Fifty-Ninth Yearbook of the National Society for the Study of Education, Part I. Chicago: University of Chicago Press, 1960.

Driver, Rosalind, Richard Gott, Sandra Johnson, Christopher Worsley, and Fiona Wylie. Assessment of Performance Unit: Science in Schools. Age 15: Report No. 1. Department of Education and Science. London, England, 1982.

Gagné, R.M., L.J. Briggs, and W.W. Wager. Principles of Instructional Design. New York: Holt, Rinehart and Winston, 1988.

Gardner, Marjorie, et al. Laboratory Assessment Builds Success. A Monograph Prepared by the Laboratory Leadership Group, Institute for Chemical Education/Berkeley, 1987-89. 1990.

Guilford, J.P. The Nature of Human Intelligence. New York: McGraw-Hill Book Company, Inc., 1967.

Hasenstab, Joe K., Walter Blood, and Eleanor Starke Kobrin. Teaching Through Learning Channels. GASC R-VIII NYSUT. Albany, NY: Performance Learning System, Inc., 1982.

Henry, Nelson B. (Ed.). Rethinking Science Education: The Fifty-Ninth Yearbook of the National Society for the Study of Education, Part I. Chicago, Illinois: The University of Chicago Press, 1960.

Lunetta, Vincent, N. and Pinchas Tamir. "Matching Lab Activities with Teaching Goals." The Science Teacher. Vol. 46, No. 5, May 1979.

Lunetta, Vincent N., Avi Hofstein, and Geoffrey Giddings. "Evaluating Science Laboratory Skills." The Science Teacher: NSTA, January, 1981.

President's Commission on Higher Education. "Higher Education for American Democracy." Vol. 1. Washington, D.C.: Government Printing Office, 1947, pp. 52.

Pett, Joel. "Fair Test EXAMINER." Winter 1990, pp. 8-9.

Phenix, H. Philip. Realms of Meaning. McGraw-Hill Book Company, 1964.

Resnick, Daniel P., and Lauren B. Resnick. "Standards, Curriculum, and Performance: A Historical and Comparative Perspective." Educational Researcher, Vol. 14, 1985, pp. 5-12.

Schofield, Beta, Patricia Murphy, Sandra Johnson, and Paul Black. Assessment of Performance Unit: Science in Schools. Age 13: Report No. 1. Department of Education and Science. London, England, 1982.

Shayer, M., D. Kuchemann, and H. Wylam. "The Distribution of Piagetian Stages of Thinking in British Middle and Secondary School Children." British Journal of Educational Psychology. February 1976, p. 46.

Skinner, S. Ballou. "Cognitive Development: A Prerequisite for Critical Thinking." Project IMPACT Training Manual. California: Orange County, Center for the Teaching of Thinking, 1983.

U.N.E.S.C.O. Unesco Handbook for Science Teachers. France: Unesco, Paris, Heinemann Educational Books, London, 1980.

Upton, Albert W. Design for Thinking. Palo Alto, California: Pacific Books, 1961.

Watson, W. and E.M. Glaser. Watson-Glaser Critical Thinking Appraisal. New York: Harcourt, Brace, and World, Inc., 1964.

Wiggins, Grant. "A True Test: Toward More Authentic and Equitable Assessment." May 1989, Phi Delta Kappan.

APPENDIX A

Illustrative Items in the Cognitive Domain

Table of Illustrative Items in the Cognitive Domain

ITEM #	ASSESSMENT CATEGORY	GRADE LEVEL	SCIENCE CONTEXT
1	Knowing	Middle School	Earth Science
2	Knowing	Middle School	Biology
3	Knowing	Middle School	Chemistry
4	Knowing	Middle School	Physics
5	Knowing	High School	Earth Science
6	Knowing	High School	Biology
7	Knowing	High School	Chemistry
8	Knowing	High School	Physics
9	Using	Middle School	Earth Science
10	Using	Middle School	Biology
11	Using	Middle School	Chemistry
12	Using	Middle School	Physics
13	Using	High School	Earth Science
14	Using	High School	Biology
15	Using	High School	Chemistry
16	Using	Middle School	Physics
17	Integrating	Middle School	Earth Science
18	Integrating	Middle School	Biology
19	Integrating	Middle School	Chemistry
20	Integrating	Middle School	Physics
21	Integrating	High School	Earth Science
22	Integrating	High School	Biology
23	Integrating	High School	Chemistry
24	Integrating	High School	Physics

Note: All items are adapted from the U. S. tests used in the Second International Science Study (Jacobson and Doran, 1989.)

ITEM #	ASSESSMENT CATEGORY	GRADE LEVEL	SCIENCE CONTEXT
1	Knowing	Middle	Earth Science

The Sun is the only body in our solar system that gives off large amounts of light and heat. Why can we see the Moon?

A. It is reflecting light from the Sun.
B. It is without an atmosphere.
C. It is a star.
D. It is the biggest object in the solar system.
E. It is nearer the Earth than the Sun.

Comments:

This item involves the two objects in our solar system with which we have the most interaction. Reflection of sunlight is involved with explanations of the phases of the moon. The item distracters represent some facts about the moon which do not explain our "seeing the Moon" (B and D) and common misconceptions about the moon (C and E). As with many good test items, one can find out more than simply how many students chose the correct response.

ITEM #	ASSESSMENT CATEGORY	GRADE LEVEL	SCIENCE CONTEXT
2	Knowing	Middle	Biology

Why are green plants important to animals?

A. Green plants consume both food and oxygen.
B. Green plants consume food and give off oxygen.
C. Green plants consume food and give off carbon dioxide.
D. Green plants produce food and give off oxygen.
E. Green plants produce food and give off carbon dioxide.

Comments:

The importance of green plants to animals and humans, in particular, is a crucial concept for an understanding of life science. Its centrality has been reaffirmed by the rebirth of environmental concerns. Most middle school science programs stress the role of green plants in producing food and giving off oxygen (choice D). The distracters represent common misconceptions about these functions.

ITEM #	ASSESSMENT CATEGORY	GRADE LEVEL	SCIENCE CONTEXT
3	Knowing	Middle	Chemistry

When butter is warmed and changes from a solid to a liquid, its molecules

A. move farther apart.
B. are reduced in number.
C. break down to simpler compounds.
D. move more slowly.
E. become smaller.

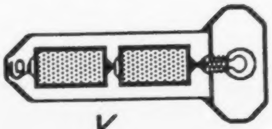
Comments:

This item assesses students' knowledge of the kinetic theory of matter. The explanation of changes of state is a key element of this theory. As matter (butter in this case) changes from a solid to a liquid, its molecules move faster and move farther apart. Therefore, choice A is correct. The distracters are common misconceptions of middle school youngsters about changes of state; changes in number or size of molecules and the chemical change to simpler compounds. Choice D above would be true for a change from liquid to solid.

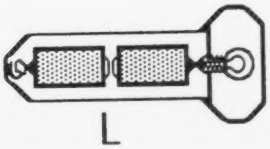
ITEM #	ASSESSMENT CATEGORY	GRADE LEVEL	SCIENCE CONTEXT
4	Knowing	Middle	Physics

A flashlight holds two batteries. In order to make it work, in which of the following ways must we place the batteries?

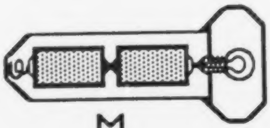
A. as in diagram K
B. as in diagram L
C. as in diagram M
D. either as in diagram L or in diagram M
E. none of these would work



K



L



M

Comments:

A pervasive activity in middle school science instruction involves series circuits consisting of batteries, bulbs, wires, and switches. In this item, students must recall that batteries in a series circuit (the flashlight) must be located in a certain way {(+) of one battery to (−) of the other} as shown in diagram K. Diagrams L and M represent logical, but wrong ways to place the batteries in the flashlight.

ITEM #	ASSESSMENT CATEGORY	GRADE LEVEL	SCIENCE CONTEXT
5	Knowing	High School	Earth Science

What binds the atmosphere to the Earth?

A. magnetic fields
B. atmospheric pressure
C. the force of gravity
D. intermolecular forces
E. wind systems

Comments:

The Earth's atmosphere is essential to life as we know it. This item tests the students' recall of the fact that gravity is the force that holds the atmosphere near the Earth's surface (choice C). "Atmospheric pressure" is the effect of the atmosphere being held near the Earth, not the cause. "Wind systems" are the result of differential atmospheric pressure on the Earth's surface. The Earth does have a magnetic field, but it does not affect the uncharged molecules of the atmosphere. The distracter "intermolecular forces" may be attractive to students who sense its "scientific sophistication."

ITEM #	ASSESSMENT CATEGORY	GRADE LEVEL	SCIENCE CONTEXT
6	Knowing	High School	Biology

Where is the energy for photosynthesis generally obtained?

A. chlorophyll
B. chloroplasts
C. sunlight
D. carbohydrates
E. carbon dioxide

Comments:

The photosynthesis process is the scientific principle that explains why green plants are so important. All the choices are involved with photosynthesis in green plants, but the energy source is sunlight (choice C). Chlorophyll is the molecule that "captures" sunlight transforming carbon dioxide and water into carbohydrate and oxygen. Chlorophyll is located in cell bodies called chloroplasts. As all the choices are involved with photosynthesis, a relatively high amount of detailed knowledge is assessed.

ITEM #	ASSESSMENT CATEGORY	GRADE LEVEL	SCIENCE CONTEXT
7	Knowing	High School	Chemistry

Which of the following particles are gained, lost, or shared during chemical changes?

A. electrons furthest from the nucleus of the atom
B. electrons closest to the nucleus of the atom
C. electrons from the nucleus of the atom
D. protons from the nucleus of the atom
E. neutrons from the nucleus of the atom

Comments:

This item assesses students' knowledge of which particles of an atom are involved with chemical changes. Almost all life processes, energy productions, and energy use are explained as chemical changes. The distracters use the three basic atomic particles; electrons, protons and neutrons. The outer electrons (farthest from the nucleus) are the ones which are gained, lost or shared as a result of chemical changes (choice A). Choice C is a choice that could identify students who are randomly guessing--as electrons are not located in the nucleus of atoms.

ITEM #	ASSESSMENT CATEGORY	GRADE LEVEL	SCIENCE CONTEXT
8	Knowing	High School	Physics

Which one of the following masses is closest to that of a small automobile?

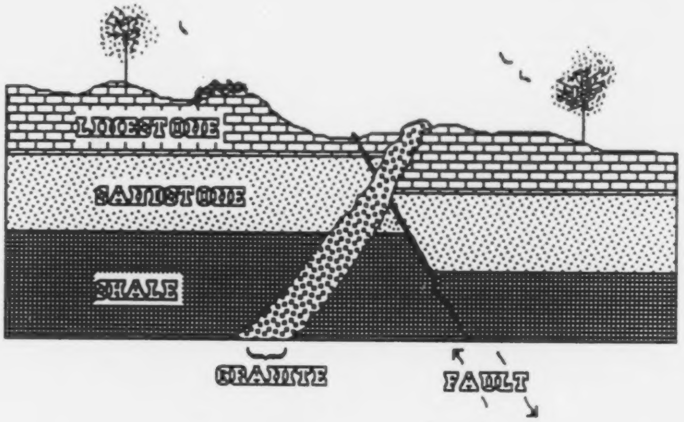
A. 100 kg
B. 1,000 kg
C. 10,000 kg
D. 100,000 kg
E. 1,000,000 kg

Comments:

This item assesses student's knowledge of one unit of the metric system--the kilogram. This unit of mass (kg) is approximately equal to 2.2 pounds. This item is an example of how to probe knowledge of this unit with a widely experienced object for teenagers--a small automobile (approximately 2,000 pounds). Therefore, choice B is correct. Choice A represents the mass of a large person (perhaps a football player) of approximately 220 pounds. Choices C, D, and E would be chosen by students who have little or no idea what the kilogram unit of mass is.

ITEM #	ASSESSMENT CATEGORY	GRADE LEVEL	SCIENCE CONTEXT
9	Using	Middle School	Earth Science

The diagram below shows a cross-section of a portion of the Earth's crust. The diagram indicates movements of the rock on one side of the fault relative to that on the other side.



Which of the rock units is the youngest?

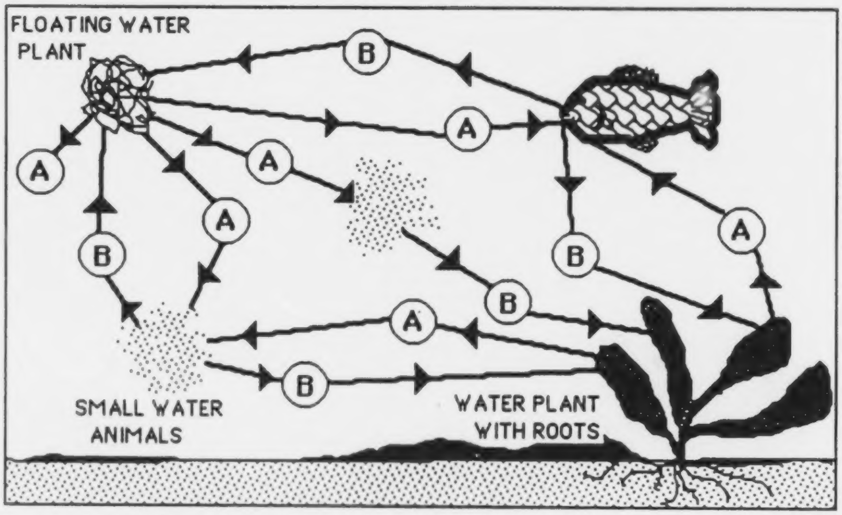
A. shale
B. granite
C. sandstone
D. limestone
E. you cannot tell from the diagram

Comments:

This item utilizes a very appropriate method of assessing the "using" cognitive level--interpretation of information from a pictorial presentation. Determining the most recent rocks is based on several principles; including "top layers are younger than lower layers" and "intrusions are more recent than layers they penetrate," etc. In this diagram, the granite intrusion penetrates all other layers and has not been overlaid or affected by the fault--so it is the most recent, making choice B correct. Limestone and shale are the top and bottom layers of the sedimentary rocks illustrated.

ITEM #	ASSESSMENT CATEGORY	GRADE LEVEL	SCIENCE CONTEXT
10	Using	Middle School	Biology

The diagram below shows a an example of interdependence among aquatic organisms. During the day the organisms either use up or give off (a) or (b) as indicated by the arrows.



Choose the right answer for (a) and (b) from the alternatives given.

A. (a) is oxygen and (b) is carbon dioxide
B. (a) is oxygen and (b) is carbohydrate
C. (a) is nitrogen and (b) is carbon dioxide
D. (a) is carbon dioxide and (b) is oxygen
E. (a) is carbon dioxide and (b) is carbohydrate

Comments:

This diagram shows the interdependence of plants and animals within an aquatic environment. Students must know that plants give off oxygen and receive carbon dioxide and correctly identify (a) as oxygen and (b) as carbon dioxide from the diagram (choice A). They also should know that animals use oxygen and give off carbon dioxide. Other distracters utilize carbohydrates--another product of photosynthesis. One distracter uses nitrogen which is the element most prevalent in air.

ITEM #	ASSESSMENT CATEGORY	GRADE LEVEL	SCIENCE CONTEXT
11	Using	Middle School	Chemistry

A dish contains 2 g (grams) of salt dissolved in 8 g of water. The dish of salty water is placed in the sunlight. 5 g of the contents of the dish evaporate into the atmosphere. What is in the remaining 5 g of solution?

A. 2 g of salt and 3 g of water
B. more than 1.5 g of salt and 3.5 g of water
C. 1 g of salt and 4 g of water
D. 5 g of water only
E. less than 0.5 g of salt and 4.5 g of water

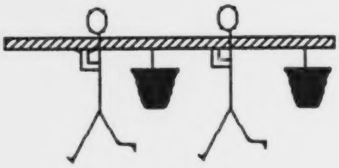
Comments:

This item tests the students' ability to apply knowledge of solutions and the process of evaporation to a very common occurrence--water evaporating in the sunlight. Students may have seen containers from which water evaporated, leaving residues or deposits. Choice A represents the correct understanding that only water evaporates leaving the 2 grams of salt--and 3 grams of water. Choice D involves the misconception that all of the salt (and some of the water) is evaporated. Choice C represents "proportional" evaporation of the salt and the water. Choices B and E involve other combinations of water and salt leaving as part of the evaporation process.

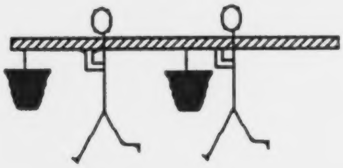
ITEM #	ASSESSMENT CATEGORY	GRADE LEVEL	SCIENCE CONTEXT
12	Using	Middle School	Physics

The diagrams below show two buckets of water being carried by two people using a pole. In which case is the load shared equally?

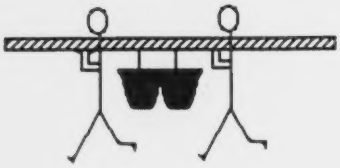
A



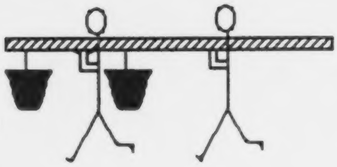
B



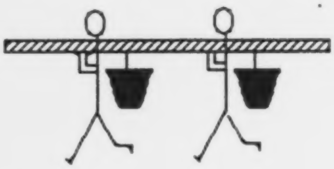
C



D



E



Comments:

The five diagrams are the choices representing five different ways in which the two people could carry the two buckets on a pole. This item requires students to apply an intuitive knowledge of forces (or weights) and torques or levers. The load is being shared equally only when both buckets are between the two people (choice C). They could also be equally shared by the two buckets being an equal distance outside each person--towards the ends of the pole.

ITEM #	ASSESSMENT CATEGORY	GRADE LEVEL	SCIENCE CONTEXT
13	Using	High School	Earth Science

The following table contains some data about different planets.

Planet	Distance from the Sun	Time for one trip around the Sun
Mercury	58 million kilometers	88 days
Venus	108 million kilometers	225 days
Earth	150 million kilometers	1 year
Jupiter	780 million kilometers	12 years
Uranus	2870 million kilometers	84 years
Neptune	4500 million kilometers	165 years

Saturn is not in the table. It is about 1430 million kilometers from the Sun.

Approximately how long will it take Saturn to make one round trip of the Sun?

A. 100 days
B. 300 days
C. 10 years
D. 30 years
E. 100 years

Comments:

An important element of understanding is interpolating within information provided. In this item, the student must "place" Saturn between Jupiter and Uranus in the table by virtue of its distance from the Sun being between that of Jupiter and Uranus. Then Saturn's time for revolution around the Sun would be between that of Jupiter (12 years) and Uranus (84 years). Choice D is the only one with a value between 12 years and 84 years. The other choices were "between" the other adjoining planets in this table. Students must recognize that the time of revolution is directly related to the distance from the Sun. The item could be answered "by memory"--so an improvement would be to use fictitious planets and/or distances and times. It also would be important to assess this "interpolating" skill with data in which the relationship is inverse.

ITEM #	ASSESSMENT CATEGORY	GRADE LEVEL	SCIENCE CONTEXT
14	Using	High School	Biology

Why is it that your body temperature does not fall even though you lose heat continually?

A. The blood distributes heat around the body.
B. Respiration results in the liberation of heat.
C. Heat is constantly being absorbed from the Sun.
D. Hot meals are eaten regularly.
E. Warm clothes are good insulators.

Comments:

This item requires that students understand the critical balance of gaining and losing heat that allows humans to maintain a constant body temperature (around 98.6°F or 37°C). The item stem states that our bodies continually lose heat--so students must identify an answer that represents (1) a gain of heat, and (2) an internal process. Choice B (cellular respiration) is the response the combines those two criteria. Choice C will only be functional when the Sun is shining--during cloudless days. Meals (choice D) provides the nutrients for cellular respiration, but cold meals provide the "cellular coal" as well as do hot meals. Warm clothes (choice E) help a body keep its heat, but it is not a supplier of heat. The blood distributes oxygen and nutrients to cells throughout the body--but does not "distribute heat."

ITEM #	ASSESSMENT CATEGORY	GRADE LEVEL	SCIENCE CONTEXT
15	Using	High School	Chemistry

When 16 g of dilute sulfuric acid is poured into 3 g of zinc in an open test tube, hydrogen gas is generated. What is the weight of the contents of the test tube after the reaction is completed?

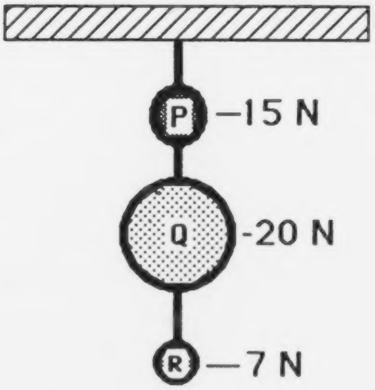
A. slightly more than 19 g
B. slightly less than 19 g
C. equal to 19 g
D. slightly less than 16 g
E. equal to 16 g

Comments:

In the chemical reaction described in this item, a gas (hydrogen) is one of the products. To answer this question correctly, students must know the principle of conservation of mass, that gases have mass, and that hydrogen gas is lighter than air (therefore it will diffuse). The weight of the solid remaining in the test tube will be "slightly less than 19 grams" (choice B). Choice A could be chosen by students who think that the hydrogen gas generated would "create" some mass. Those students who believe in the conservation of mass principle might choose choice C--equal to 19 grams, but forget that the hydrogen will escape. Choices D and E would be selected by students who might see the "slightly less" or "equal to" phrases and be attracted to that part regardless of the number of grams involved.

ITEM #	ASSESSMENT CATEGORY	GRADE LEVEL	SCIENCE CONTEXT
16	Using	High School	Physics

The objects P, Q, and R of weight 15 N (newtons), 20 N and 7 N, are hung with a light thread as shown in the figure.



What is the tension in the thread between P and Q?

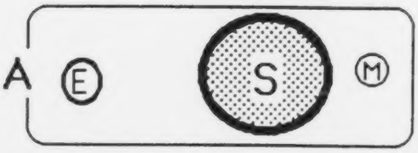
A. 42 N
B. 35 N
C. 27 N
D. 15 N
E. 7 N

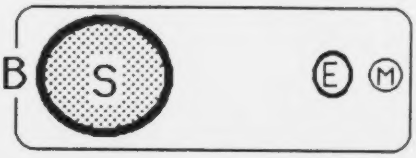
Comments:

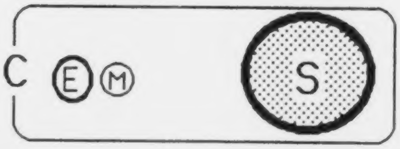
This item requires that students know that tension is the force in a thread or rope and that any segment of a supporting thread must exert enough tension (force) to hold the objects below that segment. In this diagram, objects Q and R are below segment PQ--therefore it must have a tension equal to their masses (20 Newtons plus 7 Newtons), making choice C correct. The other choices are various combinations of the three masses involved. No choice requires any concern about the mass of the light thread--so it simplifies the question. Choice A is the sum of all these masses, while choice B is the sum of the masses of objects P and Q. Choice D is the mass of the object above the segment and choice E is the mass of the object below object Q.

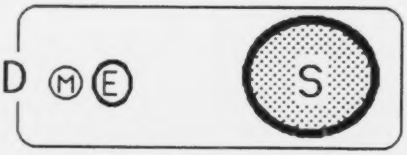
ITEM #	ASSESSMENT CATEGORY	GRADE LEVEL	SCIENCE CONTEXT
17	Integrating	Middle School	Earth Science

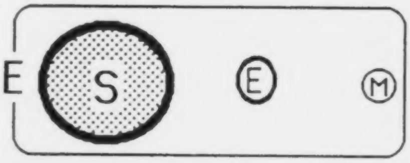
A solar eclipse occurs when the Earth, Sun and Moon are in which of the following relative positions?

A


B


C


D


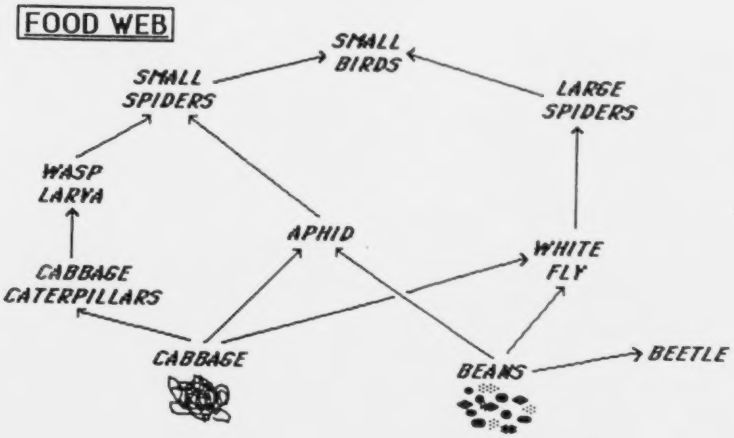
E


Comments:

Students are very familiar with the Moon, the Sun and the Earth. However, to answer this item correctly, students must connect (or integrate) several separate observations and principles. A solar eclipse occurs when people on the Earth are unable to see the Sun (during daytime). Students must realize that this can occur only when the Moon is in position to block the light from the Sun. Choice C shows the only diagram in which the Moon is between the Earth and the Sun. Choice A seems to show the Moon revolving around the Sun, instead of around the Earth. Choices B and D show configurations that could be lunar (Moon) eclipses. The diagram in choice E seems to show an inappropriate distance between the Moon, the Sun, and the Earth.

ITEM #	ASSESSMENT CATEGORY	GRADE LEVEL	SCIENCE CONTEXT
18	Integrating	Middle School	Biology

Below is a diagram showing a food web. A food web shows what the animals eat. Some animals eat the plants. These are then eaten by other animals who may be eaten by others. These are then eaten by other animals who may be eaten by others. The arrows go from the food to the eater. For example: cabbage → aphid (means cabbage is eaten by aphids)



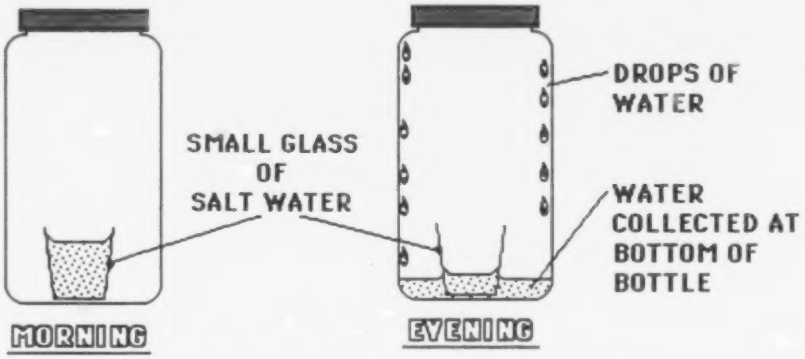
If all the beans were dug up and destroyed, which animal would disappear from this food web?

A. large spiders
B. beetles
C. aphids
D. whiteflies
E. small birds

Comments:

Understanding a food web is an important skill in biology. To answer that question, students must comprehend the meaning of the arrow (x → y) meaning that being 'x' is eaten by 'y'. Several items in the food web are eaten by several animals. An animal would disappear from this food web when its only food source is deleted from the food web. In this web, beetles were the only animals that fed exclusively on beans--the food source that was hypothesized to be destroyed, making B the correct choice. Choices C and D (aphids and whiteflies) ate beans but also could survive by eating cabbage. Large spiders and small birds are towards the top of the food web, but can survive by eating other things that do not eat beans.

ITEM #	ASSESSMENT CATEGORY	GRADE LEVEL	SCIENCE CONTEXT
19	Integrating	Middle School	Chemistry


DROPS OF WATER
SMALL GLASS OF SALT WATER
WATER COLLECTED AT BOTTOM OF BOTTLE
MORNING
EVENING

One morning a glass containing salt water is placed in a large bottle. The bottle is tightly sealed and placed in the sunlight. During the day drops of water form inside of the large bottle and collect at the bottom. No water spills out of the glass.

How salty is the water collected at the bottom of the large bottle?

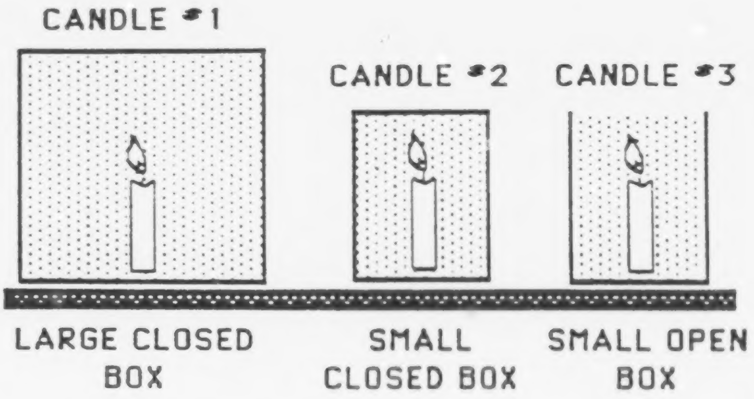
A. as salty as the water in the glass
B. a little less salty than the water in the glass
C. a lot less salty than the water in the glass
D. only a very slight taste of salt
E. no taste of salt

Comments:

This item requires students to analyze the several phenomena present in this situation and the relevant scientific principles. When the day's warmth (or sunlight) warms the salty water, some water will evaporate and then condense on the inside of the bottle as drops of water. These drops slide down the sides of the bottle and will collect at the bottom. A key principle involved is that only water molecules will evaporate, leaving the salt in the glass in the bottle. Therefore, the evaporated water collected at the bottom will have no taste of salt--choice E. The distracters represent different amounts of salt that may be present in the water at the bottom.

ITEM #	ASSESSMENT CATEGORY	GRADE LEVEL	SCIENCE CONTEXT
20	Integrating	Middle School	Physics

Three candles, which are exactly the same, are placed in different boxes as shown in the diagram. Each candle is lit at the same time.



In what order are the candle flames most likely to go out?

A. 1, 2, 3
B. 2, 1, 3
C. 2, 3, 1
D. 1, 3, 2
E. 3, 2, 1

Comments:

This item requires students to predict the order that the three flames will go out. The amount of oxygen (air) available to keep the flame burning is the factor that should be considered. Therefore, the flame in the small closed box will go first with the larger closed box next. The small open box will continue burning until the oxygen is used up, so choice B is keyed correct.

ITEM #	ASSESSMENT CATEGORY	GRADE LEVEL	SCIENCE CONTEXT
21	Integrating	High School	Earth Science

Astronomers have cataloged stars according to their color and apparent brightness. They have also been able to estimate the stars' distances from the Earth and their temperatures. Here is a table with this data for some stars.

TABLE OF OBSERVATIONS

Star	Relative Order of Brightness	Distance From Earth (Light Years)	Approximate Surface Temperature (°C)	Color
Sirius	*1*	*9*	*10,000*	*white*
Canopus	*2*	*98*	*10,000*	*white*
Arcturus	*3*	*36*	*4,000*	*reddish*
Vega	*4*	*62*	*10,000*	*white*
Aldeberan	*5*	*52*	*4,000*	*reddish*

From this data what two properties would seem most closely related?

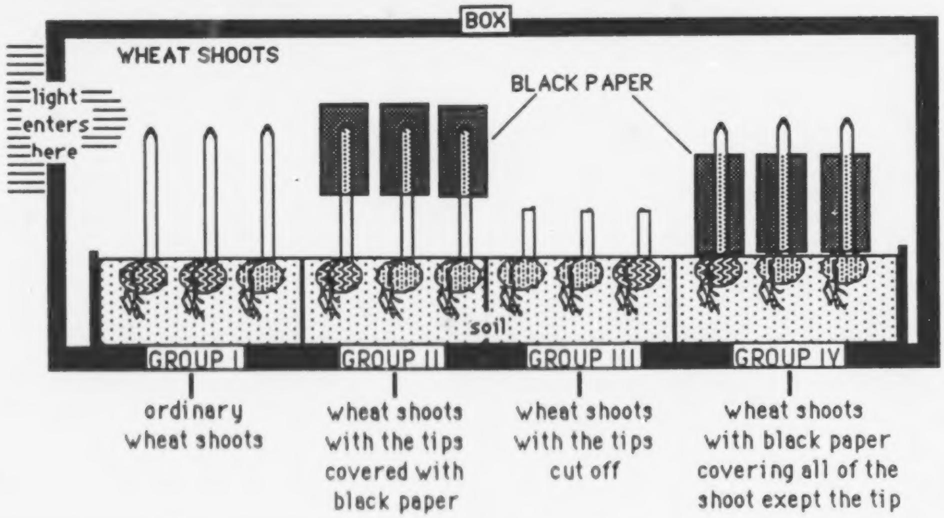
A. order of brightness and color
B. order of brightness and distance from Earth
C. color and surface temperature
D. distance from Earth and color
E. order of brightness and surface temperature

Comments:

The determination of relationships between sets of data is a key way in which scientists are able to obtain meaning from a collection of observations. Students must examine each column of data and determine what pattern exists among those data and then find which patterns match. The pattern of data on relative order of brightness is one of "constant increase." No other column of data has a similar pattern (constant increase). So any choice that involves "order of brightness" (choices A, B, and E) can be eliminated. The data on surface temperature and color do contain the same pattern: 10,000°C with white color and 4,000°C with reddish color, making choice C correct. These analytical skills are important ones for students to master.

ITEM #	ASSESSMENT CATEGORY	GRADE LEVEL	SCIENCE CONTEXT
22	Integrating	High School	Biology

The apparatus in the diagram is used to test the theory that wheat shoots grow towards light when light falls on their tips. There are four different groups of wheat shoots as indicated.


BOX
WHEAT SHOOTS
light enters here
BLACK PAPER
soil
GROUP I
GROUP II
GROUP III
GROUP IV
ordinary wheat shoots
wheat shoots with the tips covered with black paper
wheat shoots with the tips cut off
wheat shoots with black paper covering all of the shoot exept the tip

It it is true that wheat shoots grow towards light when light falls on their tips, which shoots would grow toward the light?

A. Group II
B. Group IV only
C. both Group I and Group II
D. both Group III and Group IV
E. both Group I and Group IV

Comments:

Predicting results from a well designed scientific experiment requires students to apply several scientific principles and knowledge of control variables, hypotheses, and dependent variables. In this experiment we are told the hypothesis--plants shoots grow toward the light when light falls on their tips. The diagram shows light coming only at the tip level. Group II shows wheat shoots with their tips covered. So these should not grow toward the light. Nor should the shoots in Group III which have had their tips cut off. The shoots in Groups I and IV should grow toward the light as their tips will be exposed to the light source, making choice E correct. Group IV is a critical group to the hypothesis because only the tips will be illuminated by the light.

ITEM #	ASSESSMENT CATEGORY	GRADE LEVEL	SCIENCE CONTEXT
23	Integrating	High School	Chemistry

W, X, Y, and Z are four elements. The formulae for three compounds made by combining together two of these elements are as follows:

Z_2Y
XZ
WY

What is the formula for the compound made by combining W with Z?

A. WZ
B. W_2Z
C. WZ_2
D. W_2Z_3
E. W_3Z_2

Comments:

This item is a classical example of requiring students to solve problems via a multi-step process. They must first examine the valence number of the four elements from the compounds that exist for compounds made from three pairs of the elements. From Z_2Y, students can decide if Z has a valence number of 1, than Y must have a valence number equal to 2. From XZ, students must realize that $X = 1$ if $Z = 1$, and lastly from WY, $W = 2$ as long as $Y = 2$. Therefore, to form a compound of W and Z, with $W = 2$ and $Z = 1$, the likely compound would be choice C. Choices A, D, and E represent other errors that students might make in either the first or second stage.

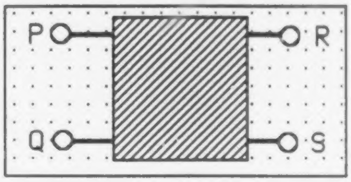
ITEM #	ASSESSMENT CATEGORY	GRADE LEVEL	SCIENCE CONTEXT
24	Integrating	High School	Physics

The figure below shows a box with four terminals: P, Q, R, and S. The following observations were made:

Observation 1 : **There is a certain amount of resistance between P and Q.**

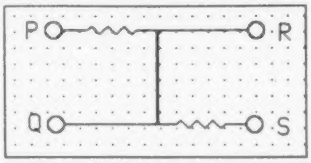
Observation 2: **Resistance between P and R is twice that between P and Q.**

Observation 3: **There is no appreciable resistance between Q and S.**

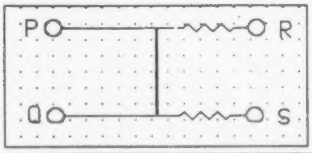


Which of the following circuits is most likely to be within the box? Assume that the resistances shown are equal.

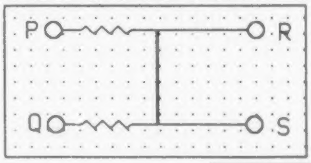
A



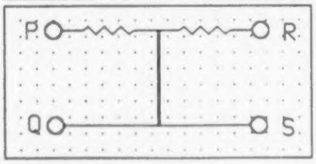
B



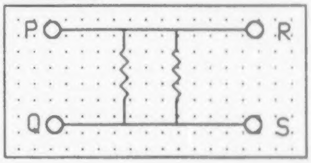
C



D



E



KEY: = wire = resistor

ITEM #	ASSESSMENT CATEGORY	GRADE LEVEL	SCIENCE CONTEXT
24	Integrating	High School	Physics

(continued)

Comments:

This item is an example of a mystery (or unknown) circuit. Students must first know what the schematic represents [straight lines = wires and zagged lines = resistors]. They must then use the three observations and mentally construct (or infer) the several pieces of the mystery circuit. Observation 3 (with no resistance between Q and S) would eliminate choices A, B, and C as they each have a resistance between terminals Q and S. Observation 2 (resistance between P and R is twice that between P and Q) is not consistent with choice E. Choice D (the correct answer) accurately portrays the resistance between P and R to be twice that between P and Q. So choice D is consistent with all observations.

APPENDIX B

Illustrative Items: Planning, Performing, and Reasoning in Scientific Problem Solving

Table of Illustrative Items: Planning, Performing, and Reasoning

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
MARGARINE	PLANNING	Identifying investigations	MIDDLE SCHOOL
BACTERIA	PLANNING	Making predictions	MIDDLE SCHOOL
IVY	PLANNING	Formulating hypotheses	MIDDLE SCHOOL
HOUSE PLANTS	PLANNING	Designing observations	MIDDLE SCHOOL
INDICATOR	PLANNING	Designing experiments	MIDDLE SCHOOL
FINDING WATER	PERFORMING	Making qualitative observations	MIDDLE SCHOOL
HOT/COLD	PERFORMING	Making quantitative observations	MIDDLE SCHOOL
MEASURE	PERFORMING	Manipulating apparatus	MIDDLE SCHOOL
CHROMATOGRAPHY	PERFORMING	Recording observations	MIDDLE SCHOOL
SINKER	PERFORMING	Performing calculations	MIDDLE SCHOOL
CABBAGE	PERFORMING	Critiquing procedures	MIDDLE SCHOOL
SLOPE	REASONING	Organizing data in tables	MIDDLE SCHOOL
BEAN HEIGHT	REASONING	Graphing data	MIDDLE SCHOOL
SOLID SORTING	REASONING	Recognizing qualitative relationships	MIDDLE SCHOOL
WOBBLE	REASONING	Recognizing quantitative relationships	MIDDLE SCHOOL
MUSHROOMS	REASONING	Recognizing limits and assumptions	MIDDLE SCHOOL
CURDLED MILK	REASONING	Explaining relationships	MIDDLE SCHOOL
ANIMAL EGGS	REASONING	Drawing conclusions	MIDDLE SCHOOL
ZAPCELL	PLANNING	Identifying investigations	HIGH SCHOOL
JUNKET	PLANNING	Making predictions	HIGH SCHOOL
SORE THROATS	PLANNING	Formulating hypotheses	HIGH SCHOOL
HEAT STORAGE	PLANNING	Designing observations	HIGH SCHOOL
BALL FALL	PLANNING	Designing experiments	HIGH SCHOOL
GRASS KEY	PERFORMING	Making qualitative observations	HIGH SCHOOL
PENDULUM	PERFORMING	Making quantitative observations	HIGH SCHOOL
READING SCALES	PERFORMING	Manipulating apparatus	HIGH SCHOOL
TESTER	PERFORMING	Recording observations	HIGH SCHOOL

Table of Illustrative Items: Planning, Performing, and Reasoning (continued)

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
SCOPE	PERFORMING	Performing calculations	HIGH SCHOOL
ROSE PETALS	PERFORMING	Critiquing procedures	HIGH SCHOOL
FLOW CHART	REASONING	Organizing data in tables	HIGH SCHOOL
LINE GRAPH	REASONING	Graphing data	HIGH SCHOOL
WIRED	REASONING	Recognizing qualitative relationships	HIGH SCHOOL
SOIL TEMPERATURE	REASONING	Recognizing quantitative relationships	HIGH SCHOOL
BICYCLES	REASONING	Recognizing limits and assumptions	HIGH SCHOOL
CARBON CYCLE	REASONING	Explaining relationships	HIGH SCHOOL
GAS BALANCE	REASONING	Drawing conclusions	HIGH SCHOOL

NOTES:

1. The item titled SCOPE was written by J.L. Accongio, 1991.
2. CHROMATOGRAPHY, SINKER, and TESTER are adapted from the Second International Science Study (Jacobson & Doran, 1989).
3. All other items are adapted from Assessment Performance Unit, Age 13: Report No. 1, and Age 15: Report No. 1, from the Department of Education and Science, London, England (1982).

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
MARGARINE	Planning	Identifying Investigations	Middle School

MARGARINE

People often say: "Margarine is better for you than butter." You cannot check whether they are right or wrong because what they say is not clear enough. Look at the suggestions below, and find the one that says what you think they mean, but says it in a way that can be checked.

Put a check mark in the box by the one you choose.

☐ A. Margarine is not so rich as butter.

☐ B. Margarine is easier to spread than butter.

☐ C. People who use margarine in cooking make better cakes.

☐ D. Margarine has a better color than butter.

☐ E. People who eat margarine live longer than those who eat butter.

DISCUSSION:

In this item, students are presented with five statements which could be substituted for the non-testable statement: "Margarine is better for you than butter." They are expected to select the statement that is not only relevant, but testable in a scientific sense.

SCORING GUIDE:

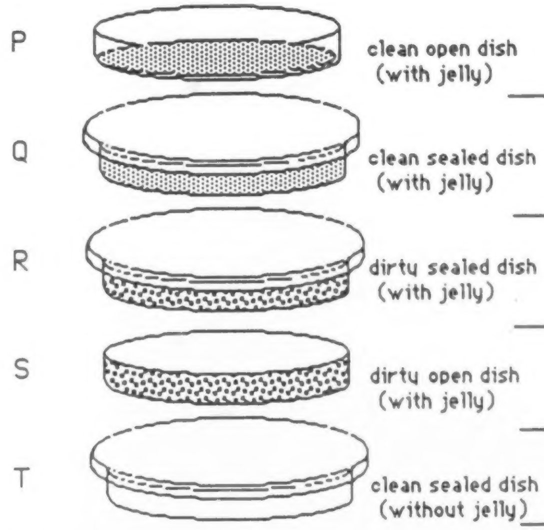
Credit is given for selecting the correct response, E. Though a retrospective or extended investigation is necessary to test the hypothesis choice E is testable and relevant, thus meeting the criteria of the question.

MAXIMUM SCORE: 1 POINT

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
BACTERIA	Planning	Making Predictions	Middle School

BACTERIA

Bacteria are present in the air. In this experiment petri dishes were used and nutrient jelly (a substance on which bacteria grow well). The clean dishes used had been sterilized to destroy any bacteria on them.



clean open dish (with jelly)	BACTERIA PRESENT
clean sealed dish (with jelly)	NO BACTERIA
dirty sealed dish (with jelly)	BACTERIA PRESENT
dirty open dish (with jelly)	
clean sealed dish (without jelly)	

What would you expect the results to be in experiments S and T? Place a check mark in the box next to the answer you choose.

- ☐ A. Both dishes would contain bacteria.
- ☐ B. Neither dish would contain bacteria.
- ☐ C. The dish in experiment S would contain bacteria but the dish in experiment T would not.
- ☐ D. The dish in experiment T would contain bacteria but the dish in experiment S would not.

DISCUSSION:

In this question the students are asked to choose a prediction based on their interpretation of written data and information in pictorial form.

SCORING GUIDE:

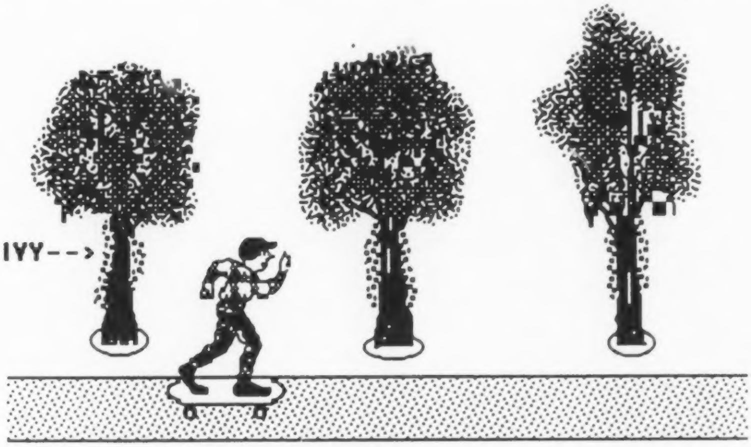
Credit is given for the selection of the correct choice C. Once the initial deduction regarding the effect of covering the dish is made there is only one possible answer.

MAXIMUM SCORE: 1 POINT

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
IVY	Planning	Formulating Hypotheses	Middle School

IVY

Skateboarding along a path, Thomas noticed that there was ivy growing on the trees but only around three-quarters of the trunks. None of the trees had ivy growing on the side nearest to the path.



Think of two different reasons why the ivy might grow only on some sides of the trees and not all around the trunks. Write them below.

Reason #1:

Reason #2:

DISCUSSION:

In this item students must apply science concepts to try to make sense of a puzzling phenomenon. They are required to suggest several alternative hypotheses based on the same set of observations.

SCORING GUIDE:

Scoring this item involves giving credit for a variety of possible hypotheses. Example responses are:

RESPONSE	POSSIBLE SCORE
• sun only shines on three sides of the trees • sun does not shine on the fourth side • because the sunlight is more intense on one side	2 points
• sun on the path side would dry up the ivy • too much sun on path side	2 points
• people walking on the path pull ivy off • people brush against the tree on the path side	2 points
• cats and dogs are more likely to have scratched off ivy on the path side	2 points
• not enough water on the path side because of shelter from wind • damper on far side	2 points
• wind blows harder on the path side--blows ivy back • caused by the wind	2 points
• may be fungus on the trees on the path side which stops ivy growing • something which attracts the ivy to one side so it won't grow on the other	1 point
• the soil is poorer on one side	1 point
• the ivy was only planted on one side	1 point
• it hasn't had time to grow on the other side yet	1 point
Reasons in terms of coincidental features, e.g., presence of path, presence of grass, presence of people, etc.	0 points
Reasons for which there is no evidence, e.g., pollution, ivy only grows on certain sides, trees are healthier on path side, etc.	0 points
Irrelevant statements, e.g., the trees branches shade the trunks, roots grow underground, birds spread the seeds, ivy is only growing on the sides of trees.	0 points
Meaningless or incomprehensible statements. Repeat of first hypothesis offered. Non-response.	0 points

MAXIMUM SCORE: 4 POINTS

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
HOUSE PLANTS	Planning	Designing Observations	Middle School

HOUSE PLANTS

A gardener buys five household plants of the same kind. She puts them side by side on a windowsill. Each day she gives them water and a little liquid fertilizer. All the plants begin to wilt. She wants to find out why they are wilting. She treats her plants as shown below.

A	B	C	D	E
water and fertilizer as before	water as before, no fertilizer	less water than before, fertilizer as before	less water than before, more fertilizer than before	no water, no fertilizer

A. Which two plants does she need to compare to find out if they wilted because she had given them too much water?

________________ and ________________

B. Which two does she need to compare to find out if the fertilizer was making them wilt?

________________ and ________________

C. Which two does she need to compare to find out if the fertilizer is helping them to grow?

________________ and ________________

DISCUSSION:

Students are asked to select the best experimental procedure for answering the questions posed. This involves knowing which variables need to be kept the same and which need changing to allow the problem to be solved.

SCORING GUIDE:

Response		Score
A.	A & C	1 point
B.	A & B	1 point
C.	C & D	1 point

MAXIMUM SCORE: 3 POINTS

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
INDICATOR	Planning	Designing Experiments	Middle School

INDICATOR

Sandra has been given 5 bottles labeled P, Q, R, S, and T with colorless liquids in them. She is told that two of the liquids are dilute acids, one is a base and the other two are water. She also has a liquid indicator (phenolphthalein). This goes colorless in an acid, red in a base, colorless in water.

Write some instructions for Sandra so that she can find out whether the liquid in each bottle is an acid, or base or just water. She is allowed to use a rack of test tubes, the indicator, and the liquids from the bottles P, Q, R, S, and T.

Make sure you describe exactly what she must do so that when she has finished she can label the bottles "acid," "base," or "water."

You may draw if you think it might help.

DISCUSSION:

In "Indicator," students are required to apply their knowledge of the behavior of a chemical indicator and suggest an investigation by means of which each of five colorless liquids could be identified as acid, base, or water.

SCORING GUIDE:

This question is to be marked by checking off columns on a grid. The checking is a way of recording how each student responds.

Noteworthy variations not allowed for by the columns on the grid should be recorded under "Teacher's Comments." The brief headings of the columns are expanded below.

Award one point each if the student indicates:

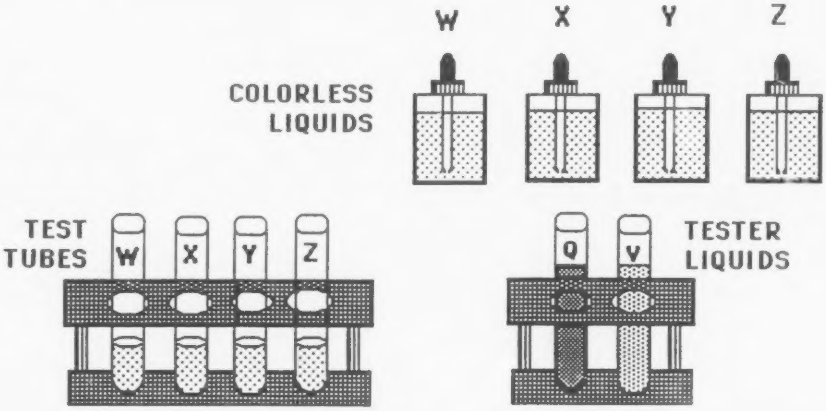
1. The five liquids are put into separate test tubes.
2. Phenolphthalein solution is added to each.
3. The liquid which turns red is the base.
4. Some other indicator should now be used to distinguish between water and acid.
5. The base identified in 3 is added to the four other unidentified liquids in test tubes (or fresh base and fresh liquids).
6. Phenolphthalein is still present by implication, or is added if needed.
7. Base should be added drop by drop to avoid "overpowering" the acid.
8. The two liquids which turn red first are water.
9. The two liquids are basic.
10. The two which remain colorless are acids.

Student No.	1	2	3	4	5	6	7	8	9	10	Teacher's Comments
	Liquids in T.T.	Add Phth.	Red = Base	Other Indic.	Add Base	Phth. Present	Drop by Drop	Red = Water	Because 2 Base	Colorless = acid	

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
FINDING WATER	Performing	Making Qualitative Observations	Middle School

FINDING WATER

You have been given four colorless liquids labeled W, X, Y, and Z. One of them is plain water. If you follow the instructions below you will be able to find out which one is water. Water is the only one that does not react with either of the two "tester" liquids labeled Q and V.



Test 1

Label four clean test tubes, as show, with the special wax pencil provided. Put a little of liquid W into tube W so that it is not more than 1/4 full. Put the other liquids into the tubes X, Y and Z, respectively. Add a few drops of the liquid called Tester Q to each tube in turn.

Write down what happens in this table.

Liquids	W	X	Y	Z
Tester Q				

(Put all your used test tubes into the container marked "USED TEST TUBES.")

(continued on next page)

Test 2

Repeat test 1, but use Tester V instead of Tester Q.

Liquids	W	X	Y	Z
Tester V				

Conclusion:

Which liquid do you think contains only water? ________________

(Put all your used test tubes into the container marked "USED TEST TUBES.")

DISCUSSION:

The purpose of this item is to assess how well students can describe the qualitative results. They are presented with four unknown liquids in dropper bottles, along with two tester liquids. The student has access to a supply of clean test tubes.

SCORING GUIDE:

Credit is earned if the answers indicate the following observations.

Test #1 Liquids	W dilute Na_2CO_3	X water	Y dilute $NaHCO_3$	Z dilute citric acid	
Tester Q (dilute HCl)	*fizz* (1 pt.)	*nothing* (1 pt.)	*fizz* (1 pt.)	*nothing* (1 pt.)	

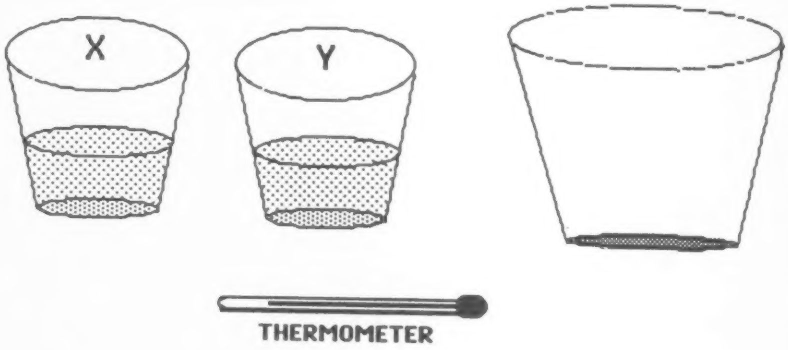
Test #2 Liquids	W dilute Na_2CO_3	X water	Y dilute $NaHCO_3$	Z dilute citric acid	
Tester V (phenolphthalein solution--slightly basic: therefore pink)	*red* (1 pt.)	*pink* (1 pt.)	*pink* (1 pt.)	*colorless* (1 pt.)	

Award 1 point if the conclusion = Liquid X

MAXIMUM SCORE: 9 POINTS

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
HOT/COLD	Performing	Making Quantitative Observations	Middle School

HOT/COLD



Use a thermometer to measure the temperatures of the water in cups marked "X" and "Y". Record the findings below.

Temperature of water in cup X = ____

Temperature of water in cup Y = ____

Pour the water from cup X and Y into the larger cup and stir. What is the approximate temperature of the mixture? ____________

DISCUSSION:

This item requires students to demonstrate their ability to make quantitative observations using a thermometer. Initial temperature of water in cup X should be 10°C, and cup Y should be 40°C.

SCORING GUIDE:

Award:
- 2 points for each reading (X, Y, and mixture) that is ± 3 degrees (1 point if ± 5 degrees)
- 1 additional point if student records appropriate unit (°C or °F) for each answer

MAXIMUM SCORE: 9 POINTS

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
MEASURE	Performing	Manipulating Apparatus	Middle School

MEASURE

This question is designed to find out how well you can measure things. (Materials: graduated ruler (in centimeters); a graduated measuring cylinder; a lever-arm balance; small beaker; filter paper; plasticene; styrofoam or plastic cup; sand; paper strips; water; envelope; wax pencil)

A. Cut off a length of paper tape 47.3 cm long. Put it in the envelope which has your name on it, and leave it at your station.

B. Measure out 55 cm^3 of water as accurately as you can and pour it into the small beaker which has your name on it. Put this to one side.

C. Put the piece of filter paper with your name on it on the balance to protect the pan. Place 68 g of plasticene on the pan. Be as accurate as you can. Leave it on the filter paper.

D. Put the cup with your name on it on the balance, and put into it 82 g of sand. Leave the sand in the cup and leave it at your station.

DISCUSSION:

This set of tasks require students to measure samples of materials using common laboratory apparatus.

SCORING GUIDE:

Each question can be awarded a total of 2 points, as shown; the score depends on how near the student's measurement is to the expected value.

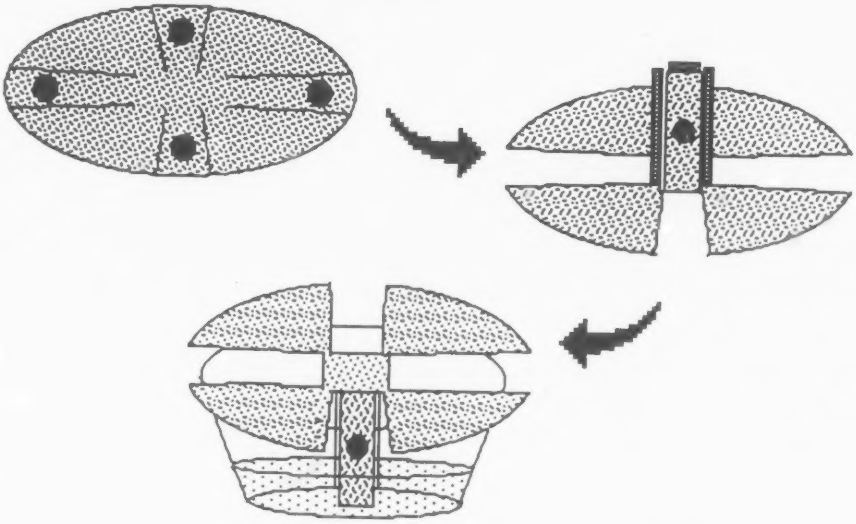
TEST	OBJECT	QUANTITY EXPECTED	CLOSE RANGE (2 pts.)	WIDE RANGE (1 pt.)	TOTAL = 8 pts. (2 pts./measurement)
a)	tape	47.3 cm.	± 0.1 cm	± 0.2 cm	
b)	water	55.0 cm^3	± 1 cm^3	± 2 cm^3	
c)	clay	68 g	± 1 g	± 2 g	
d)	sand	82 g*	± 2 g	± 4 g	

*Mass of the cup may vary by more than 1g. To check student's performance, use a container the mass of which you have already determined--or use plastic cups with known mass.

MAXIMUM SCORE: 8 POINTS

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
CHROMATOGRAPHY	Performing	Recording Observations	Middle School

CHROMATOGRAPHY



Before you are a small cup of water and a piece of cut filter paper. Bend the tabs with colored dots upward as shown in the diagram. Next, turn the paper upside-down and place the four tabs into the small cup. (BE SURE THE COLORED DOTS ARE ABOVE THE WATER SURFACE.) DO NOT LIFT THE CUP.

1. You now carefully turn the cup around on the table. Determine if the coloring from each of the dots moves at the same rate. According to what you observe, circle the correct response below.

   SAME RATE DIFFERENT RATE CAN'T TELL

   When the first color reaches the top of the paper, remove and flatten it out on a piece of toweling paper on the table.

(continued on next page)

2. Describe what happened to each of the dots.

BLACK DOT ______________________________

RED DOT ______________________________

YELLOW DOT ______________________________

GREEN DOT ______________________________

3. Give an explanation for what happened to the black dot.

____

____

____

DISCUSSION:

In this item the student is expected to follow simple directions and then record their observations. The item goes further in asking the student to offer an explanation regarding the behavior of the black dot.

Preparation of the materials for this item is essential for success. These include 1 small plastic cup (4 oz.) with water less than 1 cm deep; circular filter paper (at least 9 cm in diameter). Cut paper with four tabs (1 cm x 3 cm) as shown. Place a different colored dot 1 cm from the end of each of the four tabs. Carefully select markers that are water soluble for the black, green, and yellow dots (try highlighters or overhead transparency markers). Use a non-water soluble marker for the red dot. The distance from the edge of each dot to the end of each tab should be at least 1 cm. Each dot should be outlined with indelible ink.

SCORING GUIDE:

1. Score one point for noting that the inks move at different rates.
2. Score points for each reasonable description as in the examples below. (One point for describing movement, one point for describing separation.)

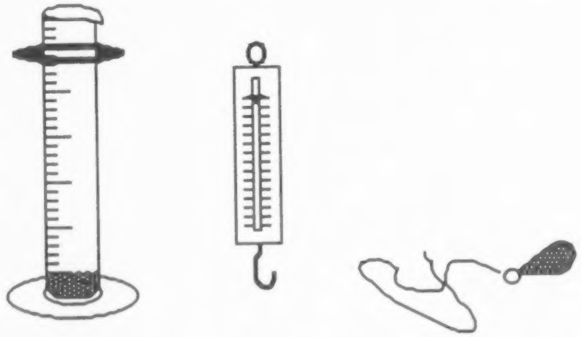
BLACK DOT:	moves up the paper, separates into several colors
RED DOT:	does not move, does not change colors
YELLOW DOT:	moves up the paper, no color separation
GREEN DOT:	moves up the paper, separates into at least two colors

3. Score 1 point for a reasonable explanation for what happens to the black dot. For example, the black ink is a mixture of many colored inks (pigments).

MAXIMUM SCORE: 10 POINTS

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
SINKER	Performing	Performing Calculations	Middle School

SINKER



Use the equipment before you to find the mass of the sinker. Then find the volume of the sinker. Now calculate the density of the sinker. Show all calculations and give the units you used to measure the mass and volume. Also give the units for density.

What is the mass of the sinker? ____________________
What is the volume of the sinker? __________________

Describe the procedure you used to find the volume of the sinker. (Show all calculations.)

What is the density of the sinker? (density = mass/volume) (Show all calculations.)

DISCUSSION:

This item requires the student to perform simple measurements to determine the volume and mass of a lead sinker. They must them perform a calculation (given the formula) for density. Materials include a 100 ml glass graduated cylinder with neck collar (clay is pressed into the bottom to prevent breakage if the lead sinker is accidentally dropped), a 0-100 g spring scale; a lead sinker (approximately 60 g) with 25 cm of string attached. (Paper towels and water should be made available.)

SCORING GUIDE:

MASS: Mass of sinker should be given. Award 2 point if mass is accurate to within .1 g and unit of measure is given; 1 point if mass is accurate to within .5 g and unit of measure is given; 0 points if unit (grams) is missing.

VOLUME: Volume of sinker should be given. Award 2 points if volume is accurate to within 1 ml and unit of measure is given; 1 point if volume is accurate within 2 ml and unit of measure is given; 0 points if unit (milliliters) is missing.

PROCEDURE FOR FINDING VOLUME: Answer should include: a) partially filled graduated cylinder; b) put sinker into graduated cylinder; c) noted initial volume and final volume (displacement); and d) subtract initial volume from final volume to determine volume of sinker. Award 1 point for each correct step listed.

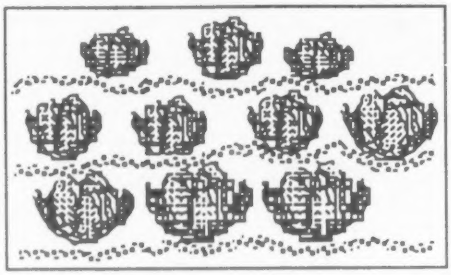
CALCULATING DENSITY: Answer should include: a) correct use of the density formula; b) calculations done correctly; c) correct units given with answer (g/ml). Award 1 point for each step done correctly.

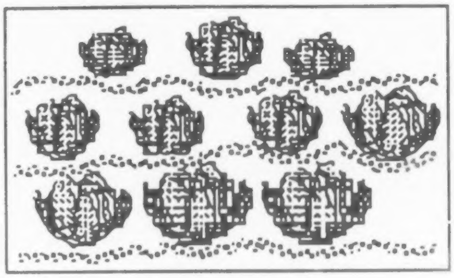
MAXIMUM SCORE: 11 POINTS

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
CABBAGE	Performing	Critiquing Procedures	Middle School

CABBAGE

A farmer is trying to find out which soil would be best for his cabbage crop. He decided to plant 10 cabbages in a section of acid soil, and 10 cabbages in basic (or alkaline) soil.


acid soil


alkaline soil

If he wants it to be a fair test he will have to make sure that some things are the same for both patches of ground.

Suggest three things that should be the same:

1. ______________________________

2. ______________________________

3. ______________________________

DISCUSSION:

This item assesses the students' ability to recognize variables that should be controlled in an experiment. The way in which the main independent variable is to be changed is indicated, but students are then asked to suggest three variables which ought to be kept constant during the investigation.

SCORING GUIDE:

Give 1 point for each different variable that would have to be controlled up to a maximum of 3.

Examples are:

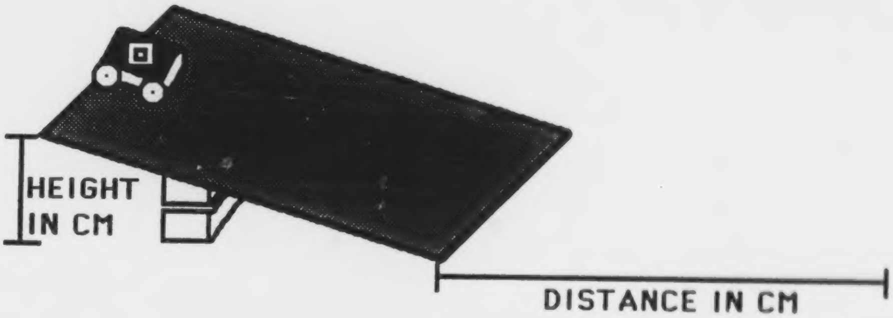
(1) Both patches get equal amounts of sunlight
(2) Both patches get equal amounts of water
(3) Both patches get equal amounts fertilizer
(4) Cabbages should be equally spaced
(5) Cabbages should all be the same variety and size
(6) Both patches should be prepared in the same way

N.B.: Each different variable given can earn a point, even if the student mentions them all in the same sentence.

MAXIMUM SCORE: 3 POINTS

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
SLOPE	Reasoning	Organizing Data in Tables	Middle School

SLOPE



Two students measured how far their truck traveled after it left the end of the board (distance "D" in cm). They changed the height of the end of the board (height "H" in cm). They put different weights in the truck. These are some of their results:

Height of raised end in cm (H)	Distance traveled in cm (D)				
	Nothing in truck	100 g in truck	200 g in truck	300 g in truck	400 g in truck
5	70	80	85		
10	85	95	103		
15	140	145	150		
20					
25					

A. They also got these results for the truck loaded with 300 g:

when H = 5 cm D = 105 cm
H = 15 cm D = 160 cm

Add these results to the table shown.

(continued on next page)

B. They put the height of the raised end up to 25 cm and found these results:

for 100 g in truck D = 195 cm
200 g in truck D = 200 cm

Add these results to the table above.

DISCUSSION:

In this item, information is presented that must be put into tabular form in "Slope." The students are required to add four numerical results from a simple investigation to a partially completed table.

SCORING GUIDE:

Height of raised end in cm (H)	Distance traveled in cm (D)				
	Nothing in truck	100 g in truck	200 g in truck	300 g in truck	400 g in truck
5	70	80	85	**105**	
10	85	95	103		
15	140	145	150	**160**	
20					
25		**195**	**200**		

A. 1 point for each correct value in the column under "300 g in truck"
B. 1 point for each correct value in the 25 cm row.

MAXIMUM SCORE: 4 POINTS

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
BEAN HEIGHT	Reasoning	Graphing Data	Middle School

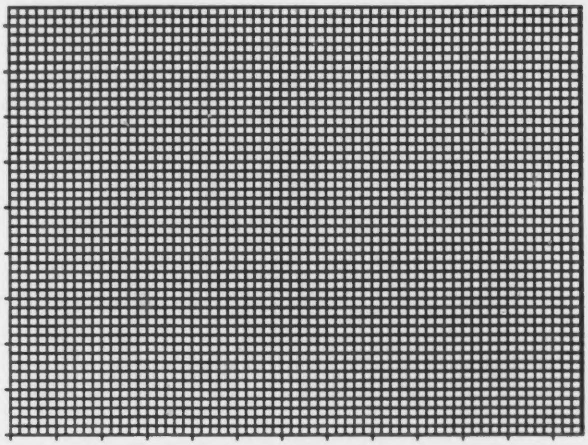
BEAN HEIGHT

Jeni measures her bean plant every week so that she could see how far it was growing. She started (0 weeks) when it was 5 cm high.

These were the heights for the first 4 weeks:

0 weeks - 5 cm
1 weeks - 15 cm
2 weeks - 30 cm
3 weeks - 40 cm
3 weeks - 45 cm

Draw and label a line graph to show how the height changed with time.



DISCUSSION:

This item asks students to represent tabular data in the form of a line graph with no help given in relation to drawing, labeling or scaling the axes.

SCORING GUIDE:

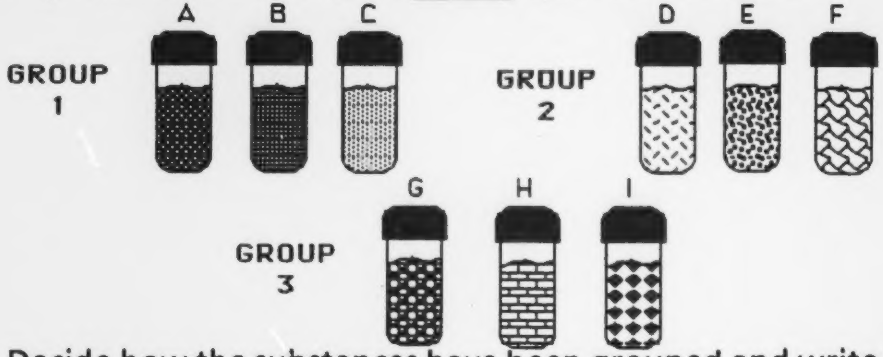
Preparation of the graph:	Score
Choice of axes (x axis = independent variable "time," and y axis = dependent variable "height")	2 points
Axes labeled *Height in cm. Time in weeks*	4 points
Suitable linear scale on x axis y axis	2 points
Plotting Data: one point to be awarded for each point correctly plotted with respect to the linear scales:	5 points
Completing smooth curve drawn: (points joined by line segments: 1 point only)	2 points

MAXIMUM SCORE: 15 POINTS

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
SOLID SORTING	Reasoning	Recognizing Qualitative Relationships	Middle School

SOLID SORTING

You have been given a collection of substances in containers from a school laboratory. The substances have been divided into three groups; 1, 2 and 3. Look carefully at the substances. You may pick them up and handle them, but do not open the containers.



A. Decide how the substances have been grouped and write this down in the spaces below.

Group 1 are all ______________________________

Group 2 are all ______________________________

Group 3 are all ______________________________

B. Now think of another way of sorting the substances into three new groups X, Y, and Z, of equal numbered containers. You may pick them up and move them around. Write down which containers are in the three groups you choose in the spaces below.

Group X are all ______________________________

Group Y are all ______________________________

Group Z are all ______________________________

PUT THE SUBSTANCES BACK INTO GROUPS 1, 2, AND 3 WHEN YOU HAVE FINISHED

DISCUSSION:

This item is designed to assess students' ability to observe similarities and differences. The tasks demand that the student identify the rule used in the classification of the 9 vials filled with different substances. Teachers must be sure to create one group containing powders, e.g., zinc dust; the next group containing irregular solids, e.g., silica gel; and the third group containing man-made regular shapes, e.g., marbles, placed in three groups of three by mass. First, the student must employ the sense of feel to determine that the vials are sorted by mass. The second sort would be based on the qualities of powders, irregular solids, and regular solids.

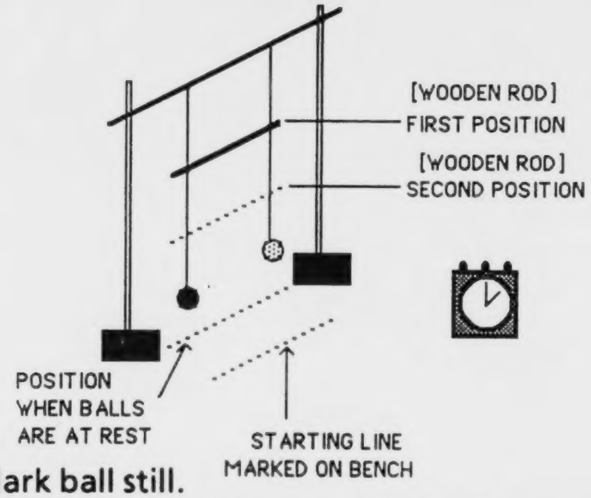
SCORING GUIDE:

One point is awarded for each correct response.

MAXIMUM SCORE: 6 POINTS

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
WOBBLE	Reasoning	Recognizing Quantitative Relationships	Middle School

WOBBLE


[WOODEN ROD]
FIRST POSITION
[WOODEN ROD]
SECOND POSITION
POSITION WHEN BALLS ARE AT REST
STARTING LINE MARKED ON BENCH

1. Hold the dark ball still.
2. Pull the light colored ball towards you to the starting line marked on the bench and let go of both balls at the same time.
3. Watch the two balls for at least one minute on the clock.
4. What do you see happening after the light colored ball is set swinging?
5. Push the wooden rod down to the second position on the string.
6. Set the light colored ball swinging again as you did before and watch both balls for one minute.
7. Describe two ways in which what you see now is different from what you saw before.

DISCUSSION:

This task involves having the students observing events and recording changes they see. This example requires that students employ the sense of sight and very little direction is given to indicate what to observe.

SCORING GUIDE:

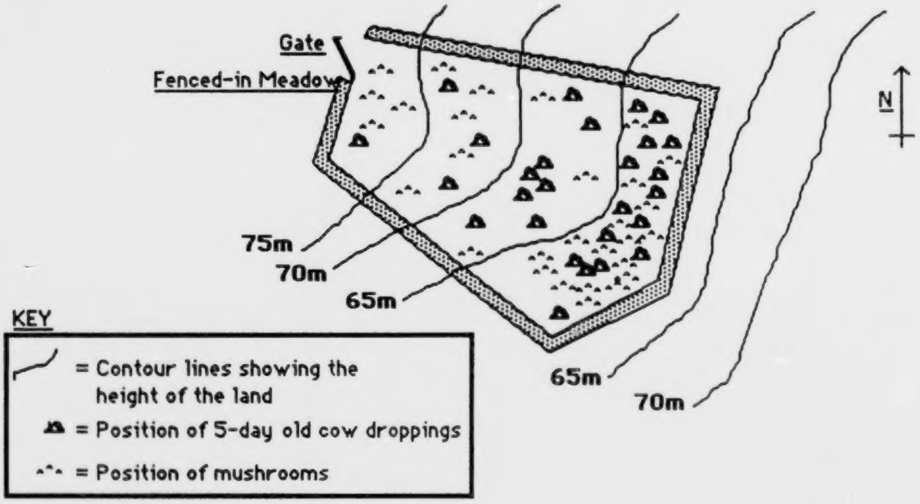
	Observations	Points
a) Position 1:	Black sphere starts moving	1
	Black sphere swings further each time	1
	Yellow sphere slows	1
	Yellow sphere (nearly) stops	1
	Wood starts to move	1
	Cycle repeats itself	1
	Black swings about as far as yellow	1
	(any 5 above)	5 points
b) Position 2:	The same changes in movement, but quicker	1
	Sideways movement less	1
	Cycle repeats itself (more rapidly)	1
	(any 2 above)	2 points

MAXIMUM SCORE: 7 POINTS

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
MUSHROOMS	Reasoning	Recognizing Limits and Assumptions	Middle School

MUSHROOMS

The map shows the distribution of mushrooms in a meadow.


Gate
Fenced-in Meadow
75m
70m
65m
65m
70m
N
KEY
= Contour lines showing the height of the land
= Position of 5-day old cow droppings
= Position of mushrooms

It is claimed that mushrooms prefer to grow in damp soil. Study the map carefully. Explain how the map either supports or does not support this claim.

DISCUSSION:

In this item students are presented with data and a generalization or hypothesis. The students must assess whether or not the hypothesis is supported by the data. Both relevant and irrelevant data are presented. In addition to supporting or rejecting the hypothesis, the students must list reasons for their judgment.

SCORING GUIDE:

Mark Scheme:

If student supports hypothesis score points if he/she mentions:

East end lower	1 point
Water will drain to it	1 point
More mushrooms appear to grow nearer the fence	1 point

If student response does not support hypothesis score points if he/she mentions:

Not enough information to decide	1 point
The soil may be well drained and no damper at any point in the field than any other	1 point
There are mushrooms growing all over the meadow	1 point

MAXIMUM SCORE: 3 POINTS

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
CURDLED MILK	Reasoning	Explaining Relationships	Middle School

CURDLED MILK

You have been given six liquids each labeled with its contents and pH value, a test tube rack with six clean test tubes, a bottle of milk and an indicator chart.

A. Place about 1 cm of milk in each test tube. To one test tube add several drops of vinegar. Shake that tube and notice what happens. Make notes below about what happens. Using a fresh test tube of milk each time, repeat this for each of the other liquids.

LIQUID	pH	WHAT HAPPENS TO MILK WHEN LIQUID IS ADDED?
Vinegar	3	
Baking Soda	9	
Lemon Juice	3	
Water	7	
Wine	5	
Caustic Soda	11	

B. Which one of the following best explains what you noticed about the liquids and the milk? The milk curdled when mixed with:

☐ A. acids only
☐ B. bases only
☐ C. bases above pH 10 only
☐ D. acids below pH 4 only
☐ E. acids below pH 4 and bases above pH 10
☐ F. both acids and bases

EMPTY THE TEST TUBES AND PUT THEM IN THE CONTAINER PROVIDED

DISCUSSION:

In this question the student has to add different solutions to milk in a test tube. After observing and recording any changes the student must then select an appropriate explanation with the aid of the pH values.

SCORING GUIDE:

		Score
a)	Vinegar - lumpy or chunky or curdled	1 point
	Baking soda - no change	1 point
	Lemon juice - lumpy or chunky or curdled	1 point
	Water - no change or gets thinner	1 point
	Wine - lumpy or chunky or curdled	1 point
	Caustic soda - no change	1 point
b)	Choice A	1 point

MAXIMUM SCORE: 7 POINTS

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
ANIMAL EGGS	Reasoning	Drawing Conclusions	Middle School

ANIMAL EGGS

ANIMAL	NUMBER OF OFFSPRING EVERY SEASON	HOW YOUNG ARE PRODUCED	AMOUNT OF PARENTAL CARE
Fish	more than 10,000	lay eggs without shells	none
Amphibian	more than 1,000	lay eggs without shells	none
Reptile	more than 100	lay eggs with shells	little
Bird	5-10	lay eggs with shells	much
Mammal	1-5	live birth	much

This table shows the number of offspring that one pair of animals belonging to these groups produces in one breeding season. It also shows how the young are produced and how much parental care is provided. (There are, of course, exceptions.)

Why do fish lay so many more eggs than reptiles, birds, and mammals?
Give two reasons using the information in the table above.

1. ____

____

2. ____

____

3. What can you conclude about the relationship between the number of young an animals produces and the amount of parenting it provides?

____

____

DISCUSSION:

This question involves the idea that predation of young organisms tends to maintain the balance of populations. Given the data in a chart, the student is expected to derive a relevant generalization.

SCORING GUIDE:

1-2. Award one point for each of the following arguments:

Eggs have no shells	1 point
More likely to be eaten	1 point
No parental care	1 point
Fewer are likely to survive	1 point
Greater number laid to maintain population	1 point

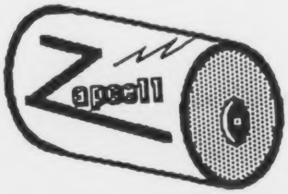
3. Award a point for any conclusion that recognizes the relationship that the greater the amount of parental care provided, the fewer number of offspring need to be produced.

MAXIMUM SCORE: 3 POINTS

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
ZAPCELL	Planning	Identifying Investigations	High School

ZAPCELL

The salesperson said: "Zapcell batteries are better."



You would not be able to check if this statement is true because it is too vague. Think about what it could mean.

Rephrase the statement and plan an experiment that would test the claim scientifically.

DISCUSSION:

In this example, students are asked to rephrase the statement, "Zapcell batteries are better." The statement is too vague or ambiguous to be tested scientifically. It is expected that students will rephrase the statement in a way that implies a workable test including the idea of controlling the test situation.

SCORING GUIDE:

Credit should be given based on the degree to which the student properly 1) rephrases the statement, 2) describes a workable test, 3) includes a mention of controls. The statement should include a description of a specific circuit with a light bulb and other materials.

MAXIMUM SCORE: 4 POINTS

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
JUNKET	Planning	Making Predictions	High School

JUNKET

A pudding called Junket is made by adding rennin to milk to make it go solid. Sheila wanted to find out the temperature at which milk with rennin solidifies the quickest. To do this, she took eight test tubes and put into each one equal quantities of milk and rennin. She put each test tube in a beaker of water at different temperatures and timed how long it took for the milk to turn solid. These are her results.

TEST TUBE	TEMP. IN °C	TIME IN MINUTES
1	10	34
2	15	23
3	20	16
4	25	12
5	30	8
6	35	---
7	40	4
8	45	4.5

Sheila dropped and broke test tube #6 so this result is missing from the table.

a) Which of the following should Sheila do in order to predict the time for milk to solidify in test tube #6? Place a check mark in the box next to the method you choose.

☐ A. Draw a pie chart
☐ B. Estimate from the numbers
☐ C. Calculate a formula
☐ D. Draw a line graph
☐ E. Use a histogram

b) Use the selected method to predict the time for milk to solidify in test tube #6. Show your work below, or submit a copy of your work.

The time predicted for milk to solidify in test tube #6 would be ______________________________

DISCUSSION:

This item challenges students to recognize a pattern in data and to identify a general trend. Here, the data are presented as a table of results in which one value is missing. Students have to determine the relationship between the variables in order to predict the missing value.

SCORING GUIDE:

a.	Student selects choice D.	1 point
b.	Student attempts to draw graph using appropriate scale	1 point
c.	Data points are plotted on graph (or reasonable attempt at other method selected)	1 point
d.	Correct answer is arrived at (5 minutes ± 30 seconds)	1 point

MAXIMUM SCORE: 4 POINTS

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
SORE THROATS	Planning	Formulating Hypotheses	High School

SORE THROATS

A doctor decided to test two new kinds of cough drops for treating sore throats. He divided the patients with the same kind of throat infection into four groups. The treatment for each group was as follows:

Group 1: **took one medicated candy cough drop (X) every 4 hours**
Group 2: **took one medicated candy cough drop (X) every 2 hours**
Group 3: **took one medicated candy cough drop (Y) every 4 hours**
Group 4: **took one ordinary candy drop (not medicated) (Z) every 4 hours**

After two days, the doctor examined the patients again. Which of these questions do you think could be answered by this test? Put a (+) mark in the box by those that could be answered. Put a (−) mark in the box by those that could not be answered. (Make sure you put a mark in every box.)

☐ A. Is a medicated candy cough drop (X or Y) a more effective treatment for a sore throat than an ordinary candy drop?
☐ B. Does it matter how often medicated candy cough drop (X) is taken?
☐ C. Does it matter how often medicated candy cough drop (Y) is taken?
☐ D. Does medicated candy cough drop (X) have a different effect from medicated candy cough drop (Y)?
☐ E. Does medicated candy cough drop (Y) more effectively treat a sore throat than an ordinary candy drop (Z)?

DISCUSSION:

Students are presented with an experimental procedure and then a series of related statements. They must relate the statements to the procedure, and mark boxes to show which statements could be tested by that experiment and which could not.

SCORING GUIDE:

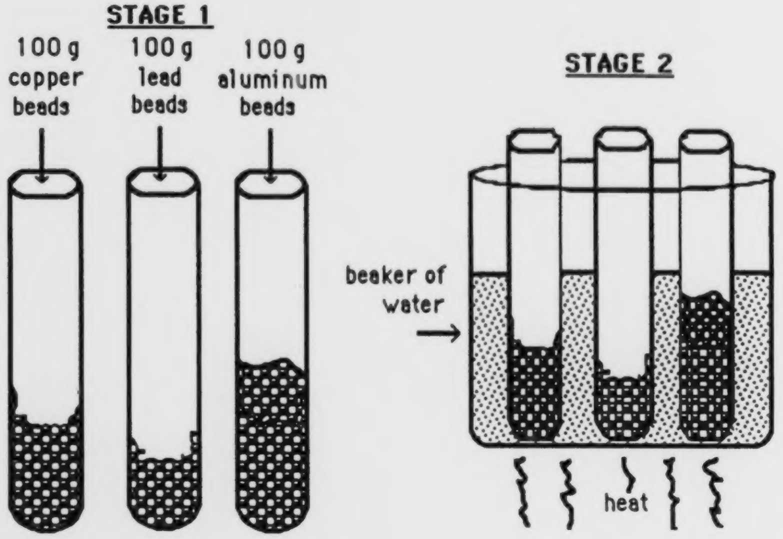
Credit is given for marking (+) next to choices A, B, D, and E, and placing a (–) next to choice C. In this format, the students are presented with a series of statements, each of which must be marked.

MAXIMUM SCORE: 5 POINTS

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
HEAT STORAGE	Planning	Designing Observations	High School

HEAT STORAGE

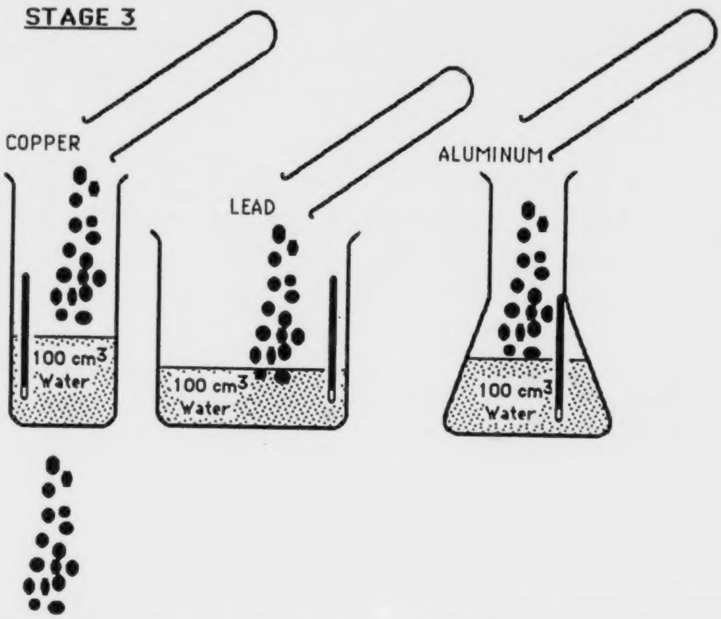
A student was doing an experiment to compare how much heat was stored in each of three different metals. This is what he did:


STAGE 1
100 g copper beads
100 g lead beads
100 g aluminum beads
STAGE 2
beaker of water
heat

He put 100g of the beads into separate test tubes.

He put the test tubes in boiling water for 10 minutes.

(continued on next page)


STAGE 3
COPPER
LEAD
ALUMINUM
100 cm3 Water
100 cm3 Water
100 cm3 Water

In STAGE 3, he tipped the contents of each test tube into separate containers, each with 100 cm³ of water and a thermometer in it. The water in all the containers was at the same temperature to start with. He then measured the temperature rise of the water in each container.

a) Using the apparatus shown, suggest two things that the student should do at STAGE 3 to make his results as accurate as possible.

b) Suggest one modification (change) to the apparatus that would make the experiment a fairer test.

DISCUSSION:

In "Heat Storage," students are asked to design a better method of observation to modify a proposed method. The students are asked to suggest changes or improvements that would make the investigation more scientific and accurate.

SCORING GUIDE:

Points should be awarded for answers such as the following:

a) 2 points for any two of the following:
- stir
- wait for maximum temperature rise
- make sure all containers were in the same ambient temperature, on the same surface, out of drafts, etc.

b) Possible modifications--award a point for any one of the following:
- use similar containers
- use lids for the containers
- use solid pieces of metal instead of beads

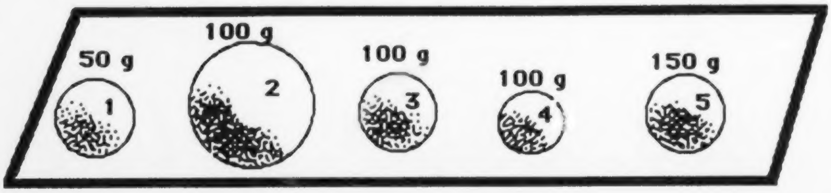
MAXIMUM SCORE: 3 POINTS

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
BALL FALL	Planning	Designing Experiments	High School

BALL FALL

You are asked to do an experiment to find out if the volume of a ball makes any difference as to the time it takes to fall to the ground.

You are given a box containing the balls shown in the diagram.



Which one of the following would you do? (Check off one box.)

☐ A. Drop balls 1, 3, and 5 from the same height and time how long each takes to fall.

☐ B. Drop ball 3 and time how long it takes to fall from different heights.

☐ C. Drop balls 2, 3, and 4 from the same height and time how long each takes to fall.

☐ D. Drop balls 1, 3, and 5 from different heights and time how long each takes to fall.

DISCUSSION:

In this question, students are asked to select the best experimental procedure for solving a given problem. This involves working out which variable needs to be kept the same and which needs changing to allow the problem to be solved. Note that drawings are important to this kind of question. It helps the student to conceptualize the problem.

SCORING GUIDE:

Students check the appropriate response. Choice C is the correct response.

MAXIMUM SCORE: 1 POINT

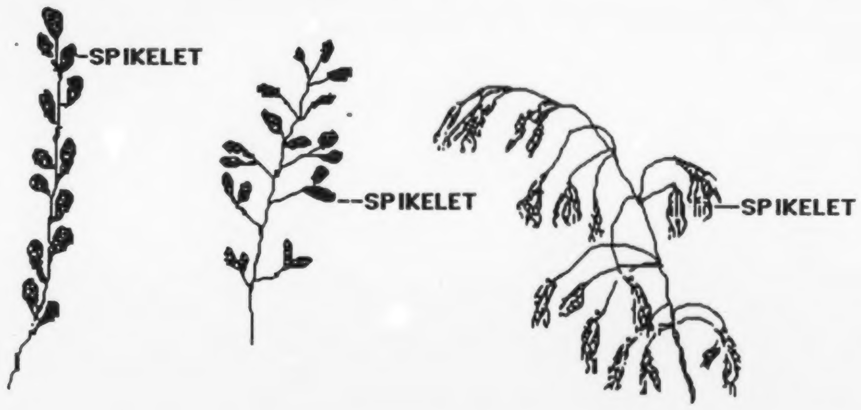
TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
GRASS KEY	Performing	Making Qualitative Observations	High School

GRASS KEY

Look carefully at the three grasses, A, B, and C, and use the identification key below to identify them.

1.	Spikelets borne on a single stem	If yes, go to 2
	Spikelets borne on branches of the stem	If yes, go to 3
2.	Spikelets clustered together and look like a single structure	If yes, go to 4
	Spikelets clearly separated	If yes, go to 5
3.	Branches bearing spikelets arise from main stem singly and each branch may have several spikelets	Meadow fescue
	Branches arise from stem in groups of two or three and each branch has only one spikelet	If yes, go to 6
4.	Spikelets form a dense cylindrical spike which is rounded at each end drooping downward	Timothy grass
	Spikelets form a dense cylindrical spike which has a point at each end	Meadow fox-tail
5.	Spikelets arranged alternately on stem	Perennial rye-grass
	Spikelets are small and clustered	Couch grass
6.	Spikelets borne on short drooping branches Lower branches are longer than upper ones giving a tapering shape	Smooth meadow grass
	Spikelets borne on long curved drooping branches and have a delicate hairy appearance	Barren brome

EXAMPLE OF SOME GRASSES



A. ______________ B. ______________ C. ______________

DISCUSSION:

This question asks students to use a classification key to identify objects or living things. They must match their observations to descriptions given in the key. The demands of reading and becoming aware of the logical structure of the key may well contribute at least as much to the student's performance as the directed observation.

SCORING GUIDE:

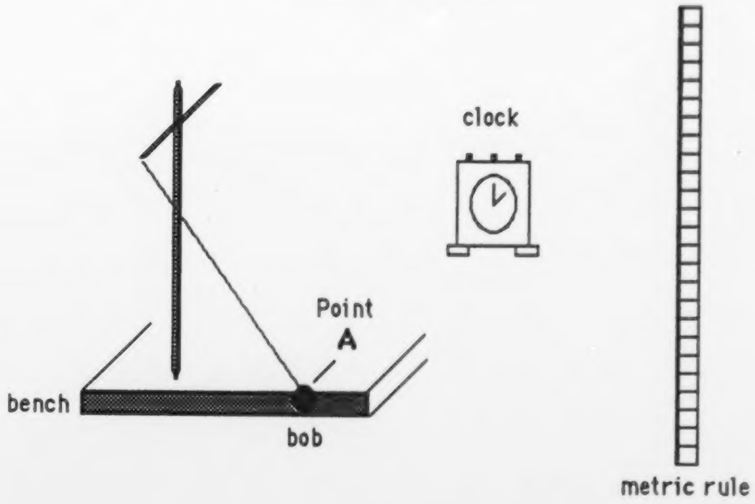
Record one point for each grass identified correctly.

A = Perennial rye grass 1 point
B = Meadow fescue 1 point
C = Barren brome 1 point

MAXIMUM SCORE: 3 POINTS

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
PENDULUM	Performing	Making Quantitative Observations	High School

PENDULUM



Think of a way of making the following measurements as accurately as possible. Find the time it takes for the bob to swing from point A and back again to the same position.

Use the space below for any calculating you want to do.

Describe how you took the measurements. ______________________

____

____

____

The length of time of one swing is ______________________

DISCUSSION:

This item requires students to demonstrate a technique for making observations of a small time interval--the period of a simple pendulum. The appropriate technique, using the apparatus provided, would be to time several swings and calculate a mean value. Great care was given in the wording of the question so that it would not suggest a technique.

SCORING GUIDE:

Award points as follows:

Student measures 1 swing only	0 points
Repeated measurements of single swing	1 point
Counted swings in a fixed period of time	2 points
Timed fixed number of swings	3 points
Approximate swing time:	
(Teacher-determined value ± 0.15 seconds)	2 points
(Teacher-determined value ± 0.25 seconds)	1 point

MAXIMUM SCORE: 5 POINTS

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
READING SCALES	Performing	Manipulating Apparatus	High School

READING SCALES

Seven different measuring instruments have been set up for you.

Do not change them in any way!

Answer as many of the questions below as you can. Don't forget to mention what units you are using. The first question has been answered to show what we mean.

What is the height of this page? *28 cm*

a. How much water is there in the measuring cylinder? ____________

b. How big is the force with which the rubber band pulls on the hook? ____________

c. What is the mass of object "X"? ____________

d. What is the temperature of the water in the flask? ____________

e. How long had the stop-clock been running before it was stopped? ____________

f. Press the push-button to switch on the current. Read the current through the circuit on the ammeter. ____________

g. Press the push-button again. Read the voltmeter. ____________

DISCUSSION:

The instruments used in this item include a thermometer, a spring scale, a voltmeter, a stop-clock, a balance, a graduated cylinder, and an ammeter. The instruments are set up and pre-set at a particular value selected by the teacher. In this way, failure by a student to score could not be due to using the instrument incorrectly. If the value lay within a predetermined range that was appropriate to the accuracy of the instrument, then the student earns 2 points. If the value given lay within a range about double that of the pre-defined range, the student earns 1 point.

SCORING GUIDE:

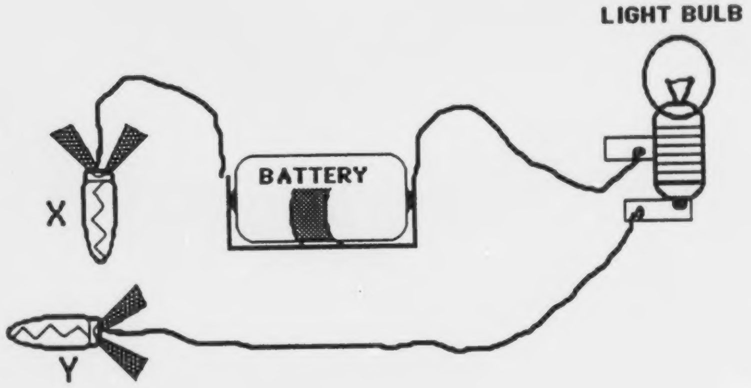
INSTRUMENT	VALUE (teacher selected)	UNITS (must be specified)	NARROW RANGE (2 points)	WIDE RANGE (1 point)
a.		cm^3	±0.5	±1
b.		N	±5	±10
c.		g	±1	±2
d.		°C	±2	±4
e.		min:sec	±0.5	±1
f.		Amp	±0.01	±0.02
g.		Volt	±0.1	±0.2

MAXIMUM SCORE: 14 POINTS

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
TESTER	Performing	Recording Observations	High School

TESTER

The diagram represents an electric tester. Using the materials provided to make an electric tester, check if an electric current flows between all the various pairs of terminals (A to F) on the circuit board. Check the tester by touching wire X to wire Y. The bulb should light. If it does not light, tell the teacher.



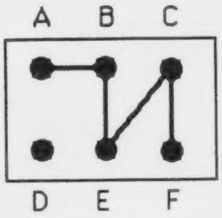
1. Perform the tests and record the results in the table below. For example, touch terminal B of the circuit board with wire X. At the same time, touch terminal C with wire Y. The bulb should light. This result has been marked for you. A positive, or plus sign (+), is in the box that is found in the column of boxes under C and in the row of boxes to the right of B. If the bulb does not light for a pair of terminals, mark a negative, or minus sign (–), in the box.

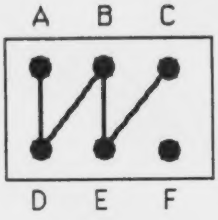
2. Use the procedure described above to get results for the other 14 ways of testing between pairs of terminals.

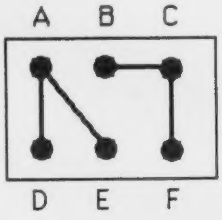
	B	C	D	E	F
A					
B		+			
C					
D					
E					

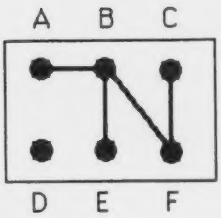
	B	C	D	E	F
A	+	+	−	+	+
B		+	−	+	+
C			−	+	+
D				−	−
E					+

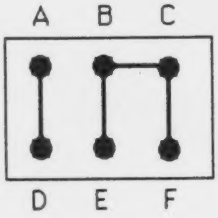
3. On the basis of the observations that you have made, which two of the following are possible ways that the terminals are connected? (Circle two.)

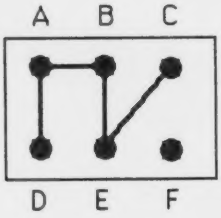
1


2


3


4


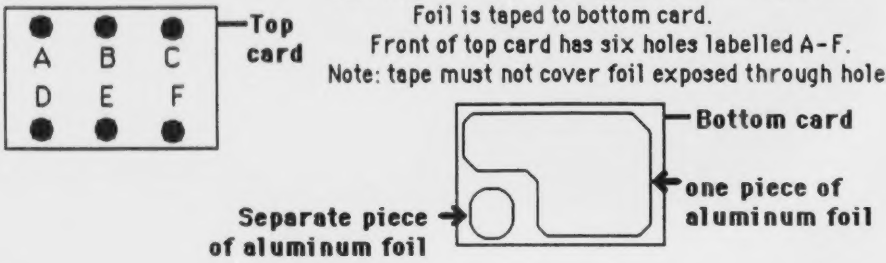
5


6


DISCUSSION:

This item requires students to make and record a series of observations then draw inferences from their findings. Construct the circuit card as follows:

CIRCUIT CARD


A
B
C
D
E
F
Top card
Bottom card
one piece of aluminum foil
Separate piece of aluminum foil

Two index cards taped together, foil in between.
Foil is taped to bottom card.
Front of top card has six holes labelled A-F.
Note: tape must not cover foil exposed through holes.

SCORING GUIDE:

The table should be completed as indicated.

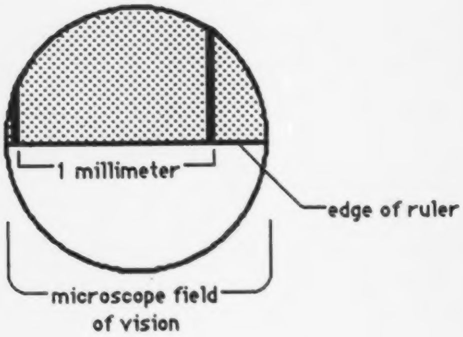
- Award 2 points if all 15 cells are marked correctly.
- Award 1 point if at least 12 cells are marked correctly.
- Award 0 points if all cells are marked " + " or all cells are marked " – ."
- Award 1 point for each correct terminal schematic selected (#1 and #4).

MAXIMUM SCORE: 4 POINTS

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
SCOPE	Performing	Performing Calculations	High School

SCOPE

A student placed the edge of a clear plastic metric ruler on the stage of a microscope. She focused on the millimeter markings so that the field of vision looked like the diagram.



She noted that the field of vision shows one millimeter and a fraction of another. (She estimated that fraction to be about 1/4 of a millimeter.) Therefore, the diameter of the field under low power was approximately 1.25 millimeters.

Using a prepared slide of paramecium, the student observed that five paramecia of the same size lined up end to end covered the diameter of the field of vision of the microscope.

Remembering that 1 millimeter = 1,000 microns, calculate the average length of the paramecia in microns. Show all calculations below:

What is the average length of the paramecia? __________

DISCUSSION:

This item requires the student to perform simple calculations to determine the length of a microscopic organism. To arrive at the correct estimation of the average length of the organisms, the student must convert the diameter of the field of vision into microns, then calculate the average by dividing by the number of organisms observed.

SCORING GUIDE:

Give one point for converting 1.25 millimeters to 1,250 microns.

Give two points for determining the average to be 250 microns per organism. (Only one point if the units are not included in the answer.)

MAXIMUM SCORE: 3 POINTS

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
ROSE PETALS	Performing	Critiquing Procedures	High School

ROSE PETALS

Bob has been asked to do a test to find out whether the coloring material in pink rose petals is a pure substance or a mixture.

He has been given this set of instructions for the test, which are not in the right order.

____ (A) Crush sand, acetone, and rose petals together using a mortar and pestle.

____ (B) Pour the liquid rose petal extract into a beaker.

____ (C) Add acetone, a drop at a time, to the middle of the colored patch on the filter paper.

____ (D) Put a few drops of rose petal extract in the middle of a filter paper.

____ (E) Put ten drops of rose petal extract into the bottle of acetone.

Decide if all the instructions listed are appropriate and then determine the right order for the test. Put 1 on the line by the side of the first thing Bob should do, 2 by the next, and so on.

DISCUSSION:

This item asks students to decide on the sequence in which the stages of an investigation must take place. In "Rose Petals," they must eliminate redundant stages; students are warned of this and asked to select the appropriate stages, then decide on their correct sequence.

SCORING GUIDE:

Award points for:

Start at A	=	1 point
Step A to B	=	1 point
Step B to D	=	1 point
Step D to C	=	1 point

Score 0 points if stage E is included anywhere in the sequence

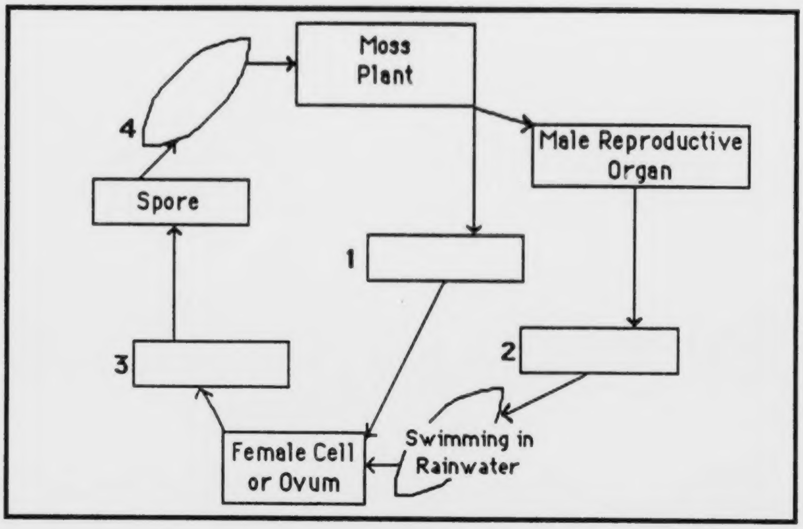
MAXIMUM SCORE: 4 POINTS

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
FLOW CHART	Reasoning	Organizing Data in Tables	High School

FLOW CHART

Read the description of the life cycle of a moss plant, and then fill in the spaces labeled 1, 2, 3, and 4 in the diagram.

> *"Mosses are small green plants which reproduce by spores produced in a capsule on a stalk. The capsule is formed from a female cell, or ovum, contained in a female reproductive organ. Before the ovum grows into a capsule, it must be joined with a male cell, or sperm, which swims in rainwater from a male reproductive organ.*
>
> *When the capsule is ripe, it opens and the spores are blown about by air currents. Each spore which lands in a suitable place splits open and grows into a new moss plant. The life cycle then begins again."*



DISCUSSION:

This item requires students to complete a flow chart by labeling the appropriate boxes according to the description of the life cycle of a moss plant.

SCORING GUIDE:

1.	Female reproductive organ	1 point
2.	Sperm (male cell)	1 point
3.	Capsule	1 point
4.	Blow in air currents (or similar)	1 point

MAXIMUM SCORE: 4 POINTS

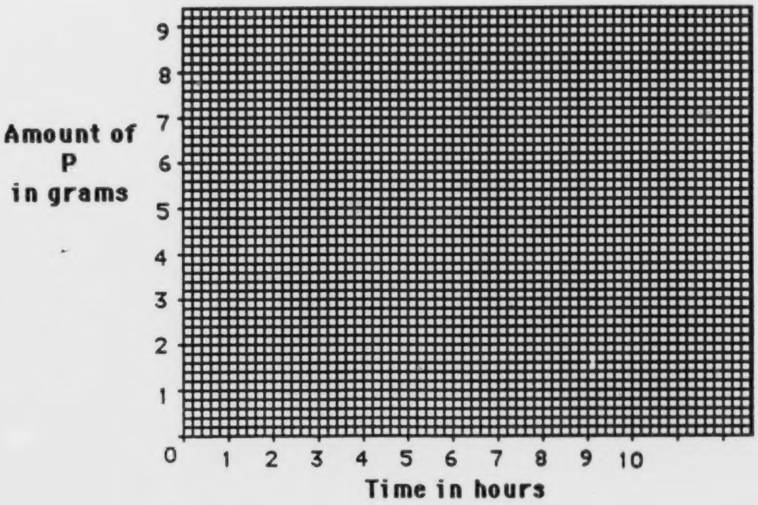
TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
LINE GRAPH	Reasoning	Graphing Data	High School

LINE GRAPH

The table shows how a substance, "P," gradually changes to another substance during a chemical reaction.

TIME IN HOURS	0	2	4	6	8	10
AMOUNT OF "P" LEFT IN GRAMS	8.0	6.4	5.3	4.4	3.7	3.2

Plot a graph to illustrate these figures.



DISCUSSION:

This question on line graphs presents data in the form of a table and asks the students to complete a graph on axes already drawn and labeled. The marking of this type of question concentrates on skill in plotting.

SCORING GUIDE:

Award 1 point for every data point plotted correctly.

Award 2 points for either a smooth curve (not necessarily going through all points or smooth curve going through all points) or straight line segments: total = 2 points

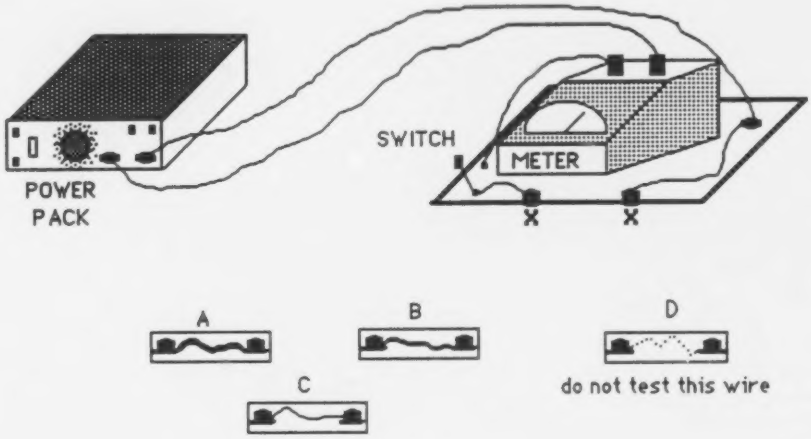
Award no points for a bar chart or for no attempt.

MAXIMUM SCORE: 8 POINTS

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
WIRED	Reasoning	Recognizing Qualitative Relationships	High School

WIRED

The circuit can be completed by connecting the apparatus labeled A, B, or C at the contact points labeled X, and then turning on the switch.



	What happened to the meter?	What happened to the wire?
• Plug A into the circuit • Switch on • Look carefully at the wire and then the meter • Switch off • Write down what you saw		

	What happened to the meter?	What happened to the wire?
• Now connect B • Switch on • Look carefully at the wire and then the meter • Switch off • Write down what you saw		
• Now connect C • Switch on • Look carefully at the wire and then the meter • Switch off • Write down what you saw		

Examine D. Do not connect it the circuit. Predict what would happen to the wire and to the meter if D was connected to the circuit.?

____

____

DISCUSSION:

This question requires the students to collect information by observation, then using what they take to be relevant and to make predictions based on them. The materials include a power source, an ammeter, and four wires mounted on carriers labeled A to D and of decreasing thickness. Students can put A, B, and C in the circuit, and have to predict what would happen to wire D if it was put into the circuit.

SCORING GUIDE:

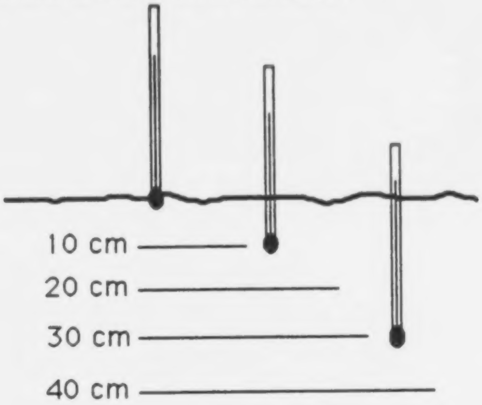
Wires A-C:	Wire A sags or expands or Wire A gets hot (red)	1 point
	Wire sags more: A → C or Wire gets hotter: A → C	2 points
Wire D:	Wire D will sag (expand)	1 point
	Wire D will get very hot or melt	1 point
Meter:	Current is measured and recorded for wire A	1 point
	Current reads lower and lower: A → C	2 points
	Current would be less than lowest reading for wire D (meter would move very little)	1 point

MAXIMUM SCORE: 9 POINTS

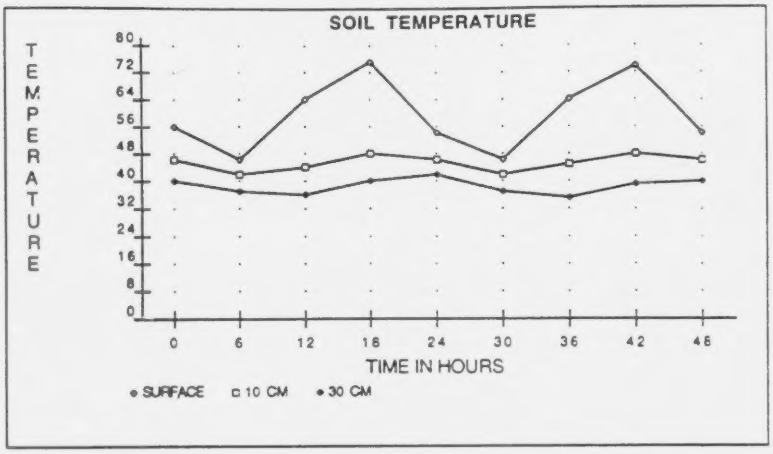
TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
SOIL TEMPERATURE	Reasoning	Recognizing Quantitative Relationships	High School

SOIL TEMPERATURE

Several scientists did an experiment to find the temperature of soil at three different depths: 1) at the surface, 2) 10 cm below the surface, and 3) 30 cm below the surface.



They measured the temperature of the soil regularly over 48 hours. The temperature at the three depths are shown in the graph below:



a) State two things the *graph* shows about the way the temperature variation depends on the depth of the soil.

1. ____

____

____

2. ____

____

____

b) About what time would the soil reach its maximum temperature at a depth of 20 cm? (Check one box below.)

☐ A. 12 hours
☐ B. 15 hours
☐ C. 18 hours
☐ D. 21 hours
☐ E. 24 hours

DISCUSSION:

Success in answering this question depends upon the student's ability to recognize and describe a pattern or relationship between two variables presented in graphical form. In addition, this item requires the student to make a prediction based on the pattern displayed.

SCORING GUIDE:

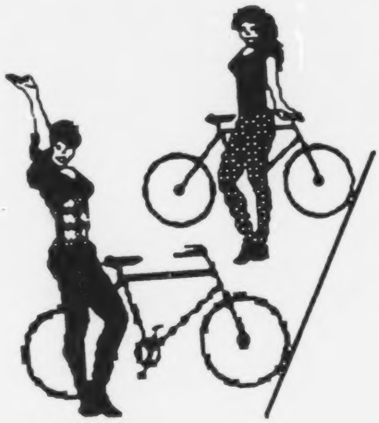
a)	1. Maxima occur at different times	1 point
	2. Difference between maximum and minimum varies with depth	1 point
b)	Choice D	1 point

MAXIMUM SCORE: 3 POINTS

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
BICYCLES	Reasoning	Seeing Limits and Assumptions	High School

BICYCLES

Stephanie and Jennifer wanted to do a test to find out which of their bicycles was the faster. This is what they did:



1. They marked out a starting line at the top of a hill.
2. They put the front wheels of their bicycles on the starting line and sat on the seats.
3. They both made five complete turns of the pedals and then sat still on their bicycles.
4. They noted which bicycle crossed the finish line first.

a) What they did was not a very fair test. Suggest two reasons for this.

1. ____

____

2. ____

____

b) Suggest one change which would make the test more fair. ____

____

____

____

____

DISCUSSION:

In this question, an experimental procedure is given, and students have to suggest reasons for poor results, propose improvements, or say why the test had been "unfair."

SCORING GUIDE:

a) Any acceptable answer that explains why it was not a fair test scores 1 point, to a maximum of 2 points. Examples:

The two girls may be different weights	1
They may push off from the ground unequally	1
They push on pedals with different forces	1
They may have had different makes of bicycle	1
The bikes may have had different gear ratios	1
One girl may be stronger than the other	1
Different shaped girls have different air resistance	1
(Any two)	2 points

b) Any acceptable answer that would make the test more fair scores 1 point. Examples:

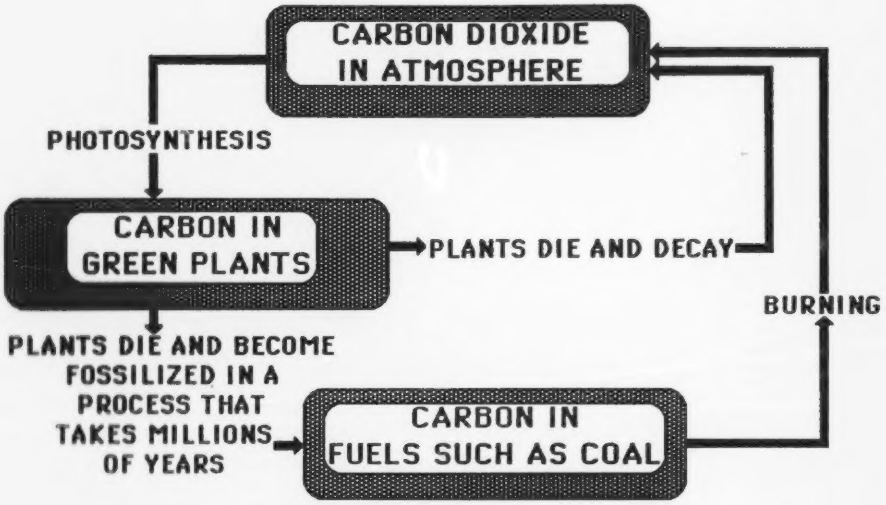
The same girl to test each bike and time the run	1
Each girl to test both bikes	1
Use weights to get the same load	1
Not to pedal--free-wheel downhill	1
Repeat the experiment	1
(Any one)	1 point

MAXIMUM SCORE: 3 POINTS

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
CARBON CYCLE	Reasoning	Explaining Relationships	High School

CARBON CYCLE

The diagram below shows part of the carbon cycle:


CARBON DIOXIDE IN ATMOSPHERE
PHOTOSYNTHESIS
CARBON IN GREEN PLANTS
PLANTS DIE AND DECAY
BURNING
PLANTS DIE AND BECOME FOSSILIZED IN A PROCESS THAT TAKES MILLIONS OF YEARS
CARBON IN FUELS SUCH AS COAL

It is said that supplies of coal will run out in about 300 years. Use the information above to explain why the coal will run out.

DISCUSSION:

In this question, students have to use information presented in a flow chart. They use the information to make a prediction based on the relationships inferred from the chart.

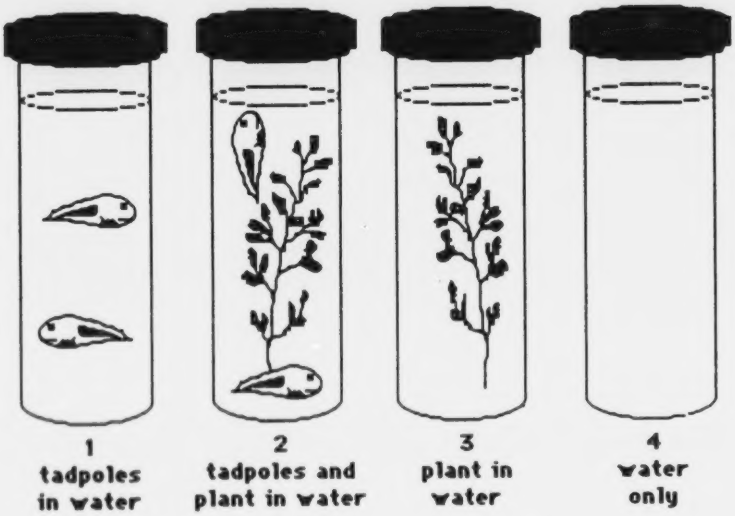
SCORING GUIDE:

Award credit for any statement indicating differences between the amount of time required for coal formation and how quickly it can be used up.

MAXIMUM SCORE: 1 POINT

TITLE	ASSESSMENT CATEGORY	SUB-CATEGORY	GRADE LEVEL
GAS BALANCE	Reasoning	Drawing Conclusions	High School

GAS BALANCE


1
tadpoles in water
2
tadpoles and plant in water
3
plant in water
4
water only

Early one morning, four containers were set up as shown in the diagram. The tops were air-tight. They were left in a well-lit place for eight hours. Describe what happened in each container.

Container #1: ____________________

Container #2: ____________________

Container #3: ____________________

Container #4: ____________________

After this time, it was suggested that the water in two of the containers would have about the same amounts of oxygen and also about the same amounts of carbon dioxide as each other.

Do you agree or disagree with the suggestion? Give your reasons.

____

____

____

DISCUSSION:

This question involves drawing conclusions based on the concept of interdependence among living things. The pertinent aspect of interdependence for this item is that oxygen, needed by animals, is produced by photosynthesis in green plants and that carbon dioxide, needed by plants, is produced by animals.

SCORING GUIDE:

Mark scheme:

Container #1:	Oxygen removed, replaced with carbon dioxide by respiration of tadpoles (1 point)*
Container #2:	Oxygen removed by tadpole and plant replaced by carbon dioxide from respiration (1 point)*
Container #3:	Carbon dioxide removed, replaced with oxygen (net result of photosynthesis and respiration in plant (1 point)*
Container #4:	Balance remains the same/no change (1 point)

* Give 1/2 point if only one gas is referred to, e.g., "oxygen produced."

If Agree:	Containers #2 and #4 contain same amount if balance is assumed (1 point)
If Disagree:	No two containers have the same amount if balance is not assumed (1 point)

MAXIMUM SCORE: 6 POINTS

APPENDIX C

Proficiency Profile for the Science Process Skills

A. PROFICIENCY PROFILE FOR THE SCIENCE PROCESS SKILL OF OBSERVING

This skill is employed in every phase of science and daily living. Observing involves using one or more of the senses in a personal interaction with something in the environment. When the five senses are insufficient, events may be observed indirectly, or with the aid of specifically developed instruments. Observations may be influenced by the background and previous experiences of the observer. When observations are made to accumulate data from which inferences will be drawn, the precision of the observation is critical. The student's level of proficiency in the skill of observing can be determined by using the following developmental profile.

Developmental Levels of Proficiency:

Level 4: Mastery

- detects rates of change
- differentiates constants from variables
- recognizes cause and effect relationships
- identifies patterns, cycles, and periodic events
- recognizes discrepancies among observations

Level 3: Competence

- repeats observations as a means to improve reliability
- uses standard measures for carefully defining observations
- orders events chronologically and sequentially over time
- identifies changes in properties

Level 2: Basic

- uses instruments to aid the senses in making observations
- refines observations to exclude false expectations
- differentiates between actual observation and personal interpretation (inference)
- distinguishes between qualitative and quantitative observations

Level 1: Minimal

- names, identifies, and distinguishes properties of objects by direct observation
    - -- seeing: color, shape, size, intensity, motion
    - -- touching: texture, shape, size, temperature, weight
    - -- hearing: pitch, volume, direction, distance
    - -- smelling: detects odors, detects differences among odors
    - -- tasting: sweet, sour, bitter, salty, metallic
- names, identifies, and distinguishes objects by using more than one sense
- moves the object being observed for different perspective to further examine its properties
- makes observations from different positions or locations

Level 0: Deficient

- skills not demonstrated or student does not possess these skills

B. PROFICIENCY PROFILE FOR THE SCIENCE PROCESS SKILL OF CLASSIFYING

Classifying involves arranging or separating objects or events according to a method, pattern, or system. The process of classifying is based on observed similarities and differences which exist between objects or events. Classification keys can be used to place items within an organized scheme, as well as to retrieve information from that scheme. The student's level of proficiency in the skill of classifying can be determined by using the following developmental profile.

Developmental Levels of Proficiency:

Level 4: Mastery

- uses an accepted classification system (key) to identify an object
- classifies data from observations as to relevancy in a particular investigation or situation
- given a set of objects, proposes several classification schemes listing different reasons for each

Level 3: Competence

- organizes data into predetermined categories
- classifies data from a series of observations to indicate frequency
- uses quantitative observation as a criterion for grouping
- creates a classification key to differentiate among a collection of similar objects
- draws simple diagram of a classification system for a group of given objects

Level 2: Basic

- groups or regroups a set of objects into subsets according to the characteristics under consideration
- groups a set of objects into subsets on the basis of two characteristics
- classifies an object using a simple flowchart, including at least three subgroups (e.g., red-blue, rectangle-circle, large-small, etc.)
- separates a set of objects into groups on the basis of likenesses and differences, and justifies groupings

Level 1: Minimal

- classifies objects according to likeness and differences by:
  - -- seeing: color, shape, size, intensity, motion
  - -- touching: texture, shape, size, temperature, weight
  - -- hearing: pitch, volume, direction, distance
  - -- smelling: detects odors, detects differences among odors
  - -- tasting: sweet, sour, bitter, salty, metallic
- groups familiar objects on the basis of use (things to wear, to play with, to eat with, to write with, etc.)
- separates a set of objects into subsets based on a single characteristic (e.g., mass, shape, or color)

Level 0: Deficient

- skills not demonstrated or student does not possess these skills

C. PROFICIENCY PROFILE FOR THE SCIENCE PROCESS SKILL OF MEASURING

All measuring activities are closely associated with science. Measuring involves developing a comparison or quantitative description of such properties as length, area, volume, weight, temperature, or pressure. Both standard and non-standard units can be used in the development of the skill. Measuring properties of objects and events measured for purposes of communication must be standardized. Measuring can be done using the metric system, as well as the English system. The student's level of proficiency in the skill of measuring can be determined by using the following developmental profile.

Developmental Levels of Proficiency:

Level 4: Mastery

- uses various techniques, such as triangulation or water displacement, to measure length, area, or volume when a tool may not be applied in direct measurement
- calibrates delicate measurement instruments
- employs optical, laser, and electronic measuring instruments

Level 3: Competence

- uses a precision of measurement relevant to the situation and instrument used
- takes readings from standard meter displays, both analog and digital
- makes prescribed settings on dials, meters, and switches
- uses formulas to generalize methods of measuring (e.g., $A = L \times W$)
- skillfully estimates measures of distance, area, volume, weight (mass), time, temperature, and speed

Level 2: Basic

- manipulates tools of measurement appropriately
- uses tools to make precise measurements
- uses approximate measures properly
- makes direct measurements of length, volume, weight, time intervals, and temperature using appropriate units
- identifies and orders objects as:
  - one-dimension - linear
  - two-dimensions - area
  - three-dimensions - volume
- can use both metric and English systems

Level 1: Minimal

- orders concrete objects by length or weight
- uses non-standard units of measure, such as pencils, pennies, thimbles, shoes, etc.
- uses standard units of measure
- selects appropriate tools for measuring various physical properties of objects

Level 0: Deficient

- skills not demonstrated or student does not possess these skills

D. PROFICIENCY PROFILE FOR THE SCIENCE PROCESS SKILL OF INFERRING

Inferring involves suggesting explanations, reasons, or causes for events which are observed. Direct observation of the event may or may not have been made. Explanations may be developed after observations of related events. Inferences, though based on judgment and evaluation, are not always valid interpretations of what really occurred. Further observations may be required to establish the validity of the inference and could lead to different inferences. Inferring requires that an individual think and reason about what has occurred, while by comparison, predicting involves suggesting what may occur in the future. The student's level of proficiency in the skill of inferring can be determined by using the following developmental profile.

Developmental Levels of Proficiency:

Level 4: Mastery

- recognizes and identifies limitations of inferences
- makes additional observations to support or validate an inference
- uses inferences to suggest need for further observations
- extends and uses inferences to formulate models

Level 3: Competence

- states several possible inferences from a group of related observation
- separates observations which support specific inferences from those which do not
- states cause-and-effect relationships from observation of related events

Level 2: Basic

- distinguishes between statements of observation and statements that are possible explanations of what is observed
- develops an inference based upon a set of previous observations
- recognizes that two people may make different inferences from the same observations

Level 1: Minimal

- makes inferences about everyday observations
- recognizes that observation is the basis for inference

Level 0: Deficient

- skills not demonstrated or student does not possess these skills

E. PROFICIENCY PROFILE FOR THE SCIENCE PROCESS SKILL OF PREDICTING

A prediction is a specific forecast of what a future observation will be. The accuracy of a prediction is strongly affected by the precision and adequacy of prior observations. Prediction is based on inference. A series of observations, and the collection and interpretation of data are important tools of prediction. An experiment can be employed to verify or contradict a prediction. The student's level of proficiency in the skill of predicting can be determined by using the following developmental profile.

Developmental Levels of Proficiency:

Level 4: Mastery

- uses quantitative measurements as a means of improving the accuracy of predictions
- uses interpolation and extrapolation as a means for making predictions
- suggests the outcome of future observations based on inferences proposed and tested
- establishes criteria for stating confidence in predictions

Level 3: Competence

- uses data from tables, charts, and graphs to make a prediction
- uses additional observations to modify predictions
- tests a prediction

Level 2: Basic

- records observations of an event over a period of time and uses this data to make a prediction
- uses a group of related events to predict an unobserved event

Level 1: Minimal

- makes predictions based upon evidence from observations of everyday occurrences
- distinguishes between guesses and predictions
- knows that predictions are based on evidence from observations

Level 0: Deficient

- skills not demonstrated or student does not possess these skills

F. PROFICIENCY PROFILE FOR THE SCIENCE PROCESS SKILL OF EXPERIMENTING

Experimenting is a way of gathering information for the purpose of testing a hypothesis. Experimenting involves understanding the use of a control and the necessity of dealing with variables. In a less formal way, experiments may be conducted simply to make systematic observations. In any experiment, there is a plan to relate cause and effect. Variables must be identified and controlled as much as possible. An experimental test of hypothesis is designed to indicate whether the hypothesis is to be supported or rejected. In designing an experiment, limitations of method and apparatus must be considered. The student's level of proficiency in the skills of designing and conducting an experiment can be determined by using the following developmental profile.

Developmental Levels of Proficiency:

Level 4: Mastery

- constructs hypotheses which can be evaluated by observation or inference
- states hypotheses in forms which suggest the variable to be manipulated
- differentiates between hypotheses which must be tested qualitatively and those which can be tested quantitatively
- states negative hypotheses in an attempt to eliminate mitigating variables
- describes the problems and limitations involved in making desired observations
- controls variables that are not part of the hypothesis being tested
- identifies sources of experimental error
- recognizes limitations of experimental apparatus
- describes limitations of experimental design
- states limitations of observations, inferences, and predictions used in answering questions or drawing conclusions

Level 3: Competence

- separates questions that can only be answered using value judgments from those that can be answered empirically
- refines questions leading to a hypothesis that can be answered by a "yes" or a "no"
- identifies constants and variables in an experiment
- makes observations and analyzes data to arrive at a conclusion
- identifies observations which are relevant to the experiment
- distinguishes between useful and extraneous data
- maintains an accurate record of experimental procedures and results

Level 2: Basic

- constructs questions that can be answered through accumulation and interpretation of data
- identifies questions leading to a hypothesis that can be answered by a "yes" or a "no"
- selects observations which assist in answering questions
- selects strategies which assist in solving problems
- conducts simple single-variable experiments and records pertinent data

Level 1: Minimal

- develops questions concerning observations, inferences, and predictions
- manipulates apparatus to make pertinent observations
- performs simple short-term activities to answer a question

Level 0: Deficient

- skills not demonstrated or student does not possess these skills

G. PROFICIENCY PROFILE FOR THE SCIENCE PROCESS SKILL OF CONSTRUCTING AND INTERPRETING MODELS

Constructing models involves the building of mental, physical, or verbal representation of an idea, object, or event as a basis for explanation and interpretations. A model is a two-dimensional or a three-dimensional image of something devised on the basis of a hypothesis. Models are used to describe and explain the interrelationships of ideas. They can be used to communicate information, demonstrate functions, mimic phenomena, or express abstract ideas. In many cases, the model implies new hypotheses; if testing these hypotheses results in new information, the model must be refined to accommodate it. The student's level of proficiency in the skills of constructing and interpreting models can be determined by using the following developmental profile.

Developmental Levels of Proficiency:

Level 4: Mastery

- explains observed phenomena by devising models
- uses measurements to develop models to scale
- extends physical or mental models to include related phenomena
- modifies existing models to include new observations
- states limitations of models
- formulates physical or mental models, idealizing observed conditions, in order to minimize variations
- devises tests for the reliability of an existing model

Level 3: Competence

- interprets models and makes inferences
- uses two or more models to explain related events
- constructs a physical representation, a drawing, or a mental image to explain observed objects or events

Level 2: Basic

- uses metaphors or analogies in describing unfamiliar things
- develops stories, drawings, charts, graphs, or diagrams to express ideas
- constructs three-dimensional models of clay, wood, paper, metal, plastic, glass, etc.

Level 1: Minimal

- recognizes certain familiar objects (such as dolls, pictures, globes, etc.) as models of real things
- makes two-dimensional models, such as drawings representing real things (machines, animals, houses, etc.)

Level 0: Deficient

- skills not demonstrated or student does not possess these skills

H. PROFICIENCY PROFILE FOR THE SCIENCE PROCESS SKILL OF QUESTIONING AND HYPOTHESIZING

Questions are problems to be solved through the application of many of the other science process skills. Precise questions can be formed on the basis of observations that have been made. The attempt to answer one question may lead to other questions, inferences, and predictions. The process of hypothesizing consists of devising a statement which can be tested by experiment. This test may either cause the hypothesis to be accepted or rejected. The student's level of proficiency in the skills of questioning and hypothesizing can be determined by using the following developmental profile.

Developmental Levels of Proficiency:

Level 4: Mastery

- constructs questions or hypotheses which can be evaluated by observation or inference
- uses controlled experiments confining a single variable to test questions or hypotheses
- states hypotheses in forms which suggest the variable to be manipulated
- differentiates between hypotheses which must be tested qualitatively and those which can be tested quantitatively
- states negative hypotheses in an attempt to eliminate variables

Level 3: Competence

- distinguishes between opinion, fact, and hypothesis
- constructs questions that can be answered though accumulation and interpretation of data
- separates questions which can only be answered using value judgments from those that can be answered empirically
- confines questions and answers to observations that have been made
- divides broad questions into parts which, when answered, will contribute to a comprehensive explanation
- refines questions leading to a hypothesis that can be answered by a "yes" or a "no"

Level 2: Basic

- states differences between questions and problems
- selects observations which assist in answering questions
- identifies questions leading to the formulation of a hypothesis that can only be answered by a "yes" or a "no"

Level 1: Minimal

- develops questions concerning observations, inferences, and predictions

Level 0: Deficient

- skills not demonstrated or student does not possess these skills

I. PROFICIENCY PROFILE FOR THE SCIENCE PROCESS SKILL OF INTERPRETING DATA

The thinking skill of interpreting data includes the sub-skills of generating, collecting, recording, organizing, analyzing, and displaying data. It requires the application of several other basic thinking skills, in particular, observing, classifying, communicating, inferring, and predicting. It is through this complex mental activity that the usefulness of data is determined in answering the question being investigated. Interpretations may be revised when additional or refined data are obtained. The student's level of proficiency in the skill of interpreting data can be determined by using the following developmental profile.

Developmental Levels of Proficiency:

Level 4: Mastery

- communicates inference or prediction with the supporting data
- selects the most acceptable interpretation from multiple interpretations of the same set of data
- restricts interpretations, inferences, explanations, and generalizations to those supported by the data

Level 3: Competence

- describes information as it is displayed on tables and graphs
- constructs bar graphs and line graphs
- interprets tables and graphs
- refines collected data to exclude irrelevant material
- distinguishes between data obtained by direct observation and interpretations of the data

Level 2: Basic

- selects data useful to answer questions under consideration
- uses collected data to make an inference or prediction
- uses diagrams, charts, and tables to organize data

Level 1: Minimal

- records simple data from observations
- describes observations in common language

Level 0: Deficient

- skills not demonstrated or student does not possess these skills

STUDENT: ______________________________ DATE: ______________

COURSE: ______________________________ GRADE: ______________

TEACHER: ______________________________

RATING SHEET
STUDENT PROFICIENCY PROFILE
SCIENCE PROCESS SKILLS

Using the set of Proficiency Profiles for Science Process Skills, rate the student on each skill by placing a check mark in the appropriate box on this table. The cells of this table are divided so you can record the pre-instruction and post-instruction proficiency profiles.

SKILL LEVEL	OBSERVING		CLASSIFYING		MEASURING		INFERRING		PREDICTING		EXPERIMENTING		CONSTRUCTING AND INTERPRETING MODELS		QUESTIONING AND HYPOTHESIZING		INTERPRETING DATA	
	PRE	POST	PRE	POST	PRE	POST	PRE	POST	PRE	POST	PRE	POST	PRE	POST	PRE	POST	PRE	POST
LEVEL 4: MASTERY																		
LEVEL 3: COMPETENCE																		
LEVEL 2: BASIC																		
LEVEL 1: MINIMAL																		
LEVEL 0: DEFICIENT																		
SKILL NOT YET ASSESSED																		

COMMENTS: (What can the student do to improve this profile?)

____

____

APPENDIX D

Illustrative Performance Tasks for Chapter 5

Table of Illustrative Performance Tasks for Chapter 5

ASSESSMENT CATEGORY	TITLE	SOURCE
Guided Investigation	Survival	APU[1]
Guided Investigation	Sugar Cubes	NAEP[2]
Guided Investigation	Volume of Rock	NAEP[2]
Guided Investigation	Water Temperature	NAEP[2]
Guided Investigation	Cold Pack	OAIP[3]
Practical Investigation	Acceleration	Rhode Island[4]
Practical Investigation	Potato Cubes	Rhode Island[4]
Practical Investigation	Acid-Base	LABS[5]
Practical Investigation	Hooke's Law	NORC-UB[6]
Practical Investigation	Identifying Gases	IPS[7]
Practical Investigation	Photosynthesis	BSCS[8]
Extended Investigation	Cost of a Shower	Connecticut[9]

[1] APU: adapted from Assessment of Performance Unit, Science in Schools, Age 15: Report No. 1. Department of Education and Science. London, England, 1982.

[2] NAEP: adapted from National Assessment of Educational Progress. Learning by Doing. Educational Testing Service, May 1987. Report No.: 17-HOS-80.

[3] OAIP: adapted from Ontario Assessment Instrument Pool. Chemistry Laboratory Achievement Test, 1979. Faculty of Education, Queen's University, Kingston, Ontario.

[4] Rhode Island: adapted from Rhode Island Distinguished Merit Program, Rhode Island Department of Education, 1990-91.

[5] LABS: Marjorie Gardner and colleagues. Laboratory Assessment Builds Success. A monograph prepared by the Laboratory Leadership Group, Institute for Chemical Education/ Berkeley, 1987-89. Copyright 1990 by the Regents of the University of California.

[6] NORC-UB: adapted from Development of Twelfth Grade Science Assessment Project. Created by the National Opinion Research Center and the University of Buffalo. Supported by NSF, 1990.

[7] IPS: adapted from Introductory Physical Science.

[8] BSCS: adapted from Biological Science Curriculum Study Newsletter article "Laboratory Tests for BSCS Students." Tamir and Glassman.

[9] Connecticut: adapted from Connecticut Common Core of Learning, Performance Assessment Project. Sponsored by the National Science Foundation. Draft: May 1, 1990.

ASSESSMENT CATEGORY	TITLE	SOURCE
Guided Investigation	Survival	APU

SURVIVAL

This task entitled "Survival" is one of the outstanding science assessment tasks developed by the Assessment of Performance Unit (APU) series in the United Kingdom during the early 1980s. The APU assessed science outcomes with students aged 11, 13, and 15 years. "Survival" was used with students aged 13 and 15 to assess the skill "performing entire investigations." This item was adapted by NAEP. The students were presented with the following challenge.

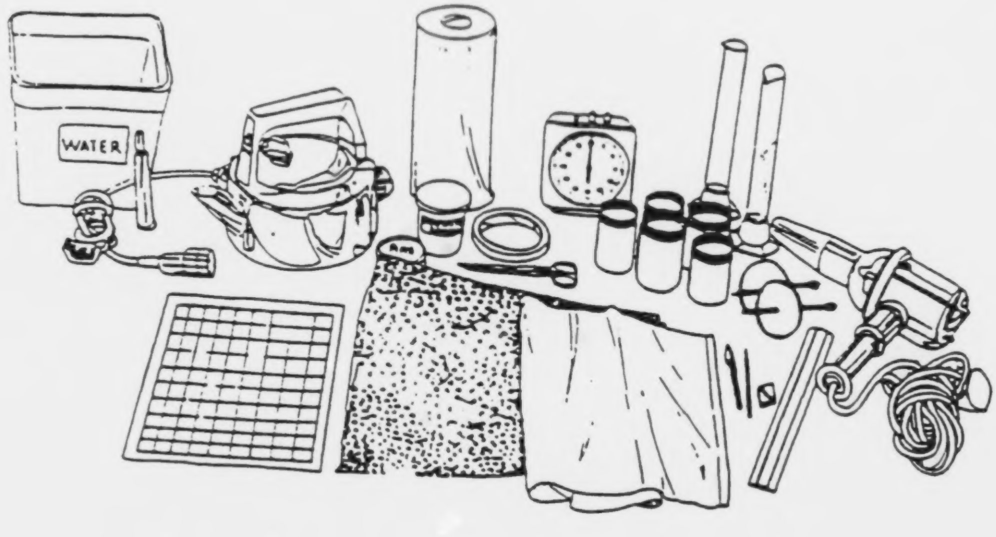
SURVIVAL

Imagine you are stranded on a mountainside in cold, dry, windy weather. You can choose a jacket made from one of the fabrics in front of you. This is what you have to find out: **Which fabric would keep you warmer?**

You can:

- use a can instead of a person.
- put warm water inside to make it more life-like.
- make it a "jacket" from the material.
- use an electric fan to make an imitation wind.

Make a clear record of your results so that another person can understand what you have found out.



(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Guided Investigation	Survival	APU

In this question, the independent variable to be tested is the type of fabric. The dependent variable to be judged is the thermal conductivity of the fabrics. The problem is set outside the laboratory situation, but the student is given clues about how to simulate the situation using laboratory apparatus.

Materials List: five cans, two having the same dimensions, aluminum (A, B), one same dimension but plastic (E), one same height but larger diameter, aluminum (C), one same diameter but shorter, aluminum (D); 2 covers for containers with holes for thermometer; thermometer; rubber bands; pins; cellophane tape; hot water source; two graduated cylinders (100 ml); sheets of fabric (cotton, wool, silk, synthetic fiber, paper, etc.); sheets of plastic for "fabric"; electric fan or hair dryer; ruler (30 cm); graph paper; cold water in container; paper towel; pencils, pens; stopwatch.

Notes for Administrator: Make available pencils, pens, erasers; stopwatch. Put out sufficient fabric for each test situation, i.e., three sheets of each kind. Remember to provide stopclock, etc. Heat the water prior to testing.

The APU staff developed a checklist for the test administrator. It is organized to follow the likely sequence of student activities. The checklist is included here. The administrator merely had to check actions that students did, leaving blank those which referred to activities that were not done by the students.

SURVIVAL	STUDENT NUMBER									
	1	2	3	4	5	6	7	8	9	. . .
Material used - blanket										
- plastic										
- thermometer										
- can A										
B										
C										
D										
E										
Used material for 1 layer										
> 1 layer										
- to cover base										
Material fixed in place										
Material held in place										
Material used alone										
Used hot water (>60°C) in can										

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Guided Investigation	Survival	APU

SURVIVAL	STUDENT NUMBER									
	1	2	3	4	5	6	7	8	9	. . .
Used warm water (35°-60°C) in can										
Used cold water (<35°C) in can										
Water measured by cylinder										
Water measured by eye										
Actual volume (ml)										
Read init. temp. of water before cooling period										
Actual temperature (°C)										
Started clock										
- within ± 5 sec. of reading temp.										
Records temp. at regular intervals										
Number of records: 2										
>2										
Read final temp. of water										
Read final temp. after set temp. drop										
Time interval < 2 minutes										
> 5 minutes										
Actual time (sec.)										
Actual final temp. (°C)										
Used electric fan - alternating between cans										
- directed between cans										
- on one material only										
Distance from material										
- < 5 cm										
- > 5 cm										
Number of trial(s) used										

ASSESSMENT CATEGORY	TITLE	SOURCE
Guided Investigation	Sugar Cubes	NAEP

SUGAR CUBES

The "Sugar Cubes" task was used in the NAEP pilot project with students at the third and seventh grades to assess their skill in "Conducting a Complete Experiment." NAEP administrators used prepared scripts to present complete experiments to individual students. Most of the scripts contained brief background information on the problem itself, and an explanation of the equipment available to investigate it. As each student worked, his/her activities were recorded by the administrator on a detailed checklist covering students' approaches to the problem, including how they set up the experiment, manipulated the variables, and measured the outcome. The administrator encouraged students to make notes and record findings on a response sheet.

SUGAR CUBES

Students are given laboratory equipment and asked to determine which type of sugar--granulated or cubed--dissolves faster in warm water, and to determine whether stirring or not stirring the mixture makes a difference. To complete this investigation, students need to identify the variables to be manipulated, controlled, and measured. They also need to make reliable and accurate measurements, record their findings, and draw conclusions.

Equipment required: six small glass beakers; sugar cubes; six packages of granulated sugar with same mass of sugar as in one cube; paper towels; a measuring cup; hot water in thermos (50-60°C); two stirrers; a timer; a graduated beaker; a graduated cylinder; a small ruler; paper; pencil.

Students are asked to do the following:

1. Briefly describe what you did to compare how fast the two kinds of sugar dissolved.
2. Write your explanation of these results.

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Guided Investigation	Sugar Cubes	NAEP

SCORING GUIDE:

NAEP administrators recorded student strategies for determining--with accurate and reliable measurements--whether loose sugar or sugar cubes dissolved at a faster rate. Successful strategies included:

1. Strategies used to compare how fast the two kinds of sugar dissolved:

   - tested both types of sugar; and
   - tested each by stirring and not stirring; and
   - maintained equal and/or consistent rates when stirring; and
   - measured to ensure equal temperatures and equal amounts of sugar and equal amounts of water for each test.

   (Score 1 point for each.)

2. Student's explanation of results:

   Score 3 points: if student notes that amount of water, rate of stirring, and temperature need to be controlled.

   Score 2 points: if student only notes one of the variables.

   Score 1 point: if student simply described procedure used or what was observed.

ASSESSMENT CATEGORY	TITLE	SOURCE
Guided Investigation	Volume of Rock	NAEP

VOLUME OF ROCK

As part of the 1972-73 NAEP survey of science, a set of performance tasks were administered to a sample students of age 9, 13, and 17 years. They were administered individually to students by a test administrator.

The "Volume of Rock" activity was administered to 13- and 17-year old students to assess their skill with a simple testing procedure, determining the volume of an irregular solid. The students were asked to determine the volume of a small rock. Each student was given a nonporous rock, a 12-inch ruler, a graduated cylinder, spring scales, water in a jar, a piece of string, and a set of instructions.

VOLUME OF ROCK

The observer reads the following instructions to the students:

> *"In front of you are a small rock and several pieces of apparatus. You are to use whatever apparatus you need to find the* ***VOLUME*** *of the small rock. List all procedures and record all measurements you make in* ***Part A*** *below. I will be making the same measurements in the same way that you do. When you have determined the volume of the rock, record your answer in* ***Part B****."*

STUDENT RESPONSE SHEET

Part A:

Procedures Used	Measurements
____________________	____________________
____________________	____________________
____________________	____________________
____________________	____________________

Part B:

The volume of the rock is ____________________.
Use the space below to do any calculations or other work.

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Guided Investigation	Volume of Rock	NAEP

ADMINISTRATOR'S GUIDE SHEET

Probe:
If student does not proceed, say, *"Think of some measurements you could make which would give you the volume of the rock."* Observer records student's actions in the following checklist.

A. Indicate the equipment the student uses.

Yes	No	
( )	( )	Graduated cylinder and water
( )	( )	Graduated cylinder and no water
( )	( )	Ruler
( )	( )	Spring scales
( )	( )	String

B. Indicate each measurement student makes. Check each measurement before student continues and record YOUR readings below.

Yes	No	
( )	( )	Initial water volume ____________ ml
( )	( )	Final water volume ____________ ml
( )	( )	Other (specify):

C. Did student calculate final volume minus initial volume?

( ) Yes
( ) No

In section A remember to indicate all the equipment used by student.

D. Additional observations:

ASSESSMENT CATEGORY	TITLE	SOURCE
Guided Investigation	Water Temperature	NAEP

WATER TEMPERATURE

The following item was also from the 1972-73 NAEP survey. The "Water Temperature" activity was administered to students aged 9, 13, and 17 years.

WATER TEMPERATURE

Administrator's Script:

SAY: *"In front of you are three styrofoam cups, two plastic measuring cups, two Centigrade thermometers, a plastic bottle containing cold water* (the water in the bottle should be as cold as is readily available), *a thermos containing hot water, and a few paper towels.*

DO: Pour 50 ml of the cold water into one of the two measuring cups.

SAY: *"Now, I am measuring out 50 milliliters of cold water into one of these two measuring cups labeled 'cold.'"*

"Now, I will pour hot water into one of your cups labeled 'hot.'"

"The two cups are for you to work with. Remember, the one labeled "HOT" contains an amount of hot water and the one labeled "COLD" contains the same amount of cold water."

DO: Give student the two cups.

SAY: *"I have also given you two thermometers. Do you know how to use a thermometer?"*

If the response is "Yes," go to II; if "No," "I don't know," or no response, go to I.

I: Demonstrate how to use a thermometer. Include: how to hold it carefully, insert it in water so as not to tip the cup, wait a period of time, and read it. (Go to II.)

II: **SAY:** *"Put a thermometer into each cup of water and record the temperatures on* ***Parts A and B*** *of your response sheet."*

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Guided Investigation	Water Temperature	NAEP

Student Response Sheet:
Complete the following sentences:

Part A. The temperature of the hot water is ___________ °C.
Part B. The temperature of the cold water is ______________ °C.
Part C. I THINK the temperature of the mixture will be ______________ °C.
Part D. The temperature of the mixture is _____________ °C.

- -

Directions to the Administrator:

After the student reads and records the temperatures of the hot and cold water on **Parts A and B** of the response sheet, read and record your findings below. If the discrepancy between your readings and the student's readings exceeds two degrees, help student to read the thermometers and to record correct readings.

___________ °C (hot water) ___________ °C (cold water)

Did you help the student read the thermometers. ( ) Yes ( ) No

DO: Place the third styrofoam cup in front of the student.

SAY: *"We are going to mix the hot and cold water together in this cup. Tell me in words what the temperature of the mixture will be compared to the temperatures of the hot and cold water."*

DO: Record the student response. _

_ _

Probe: If the student:

a. gives a numerical response, ask, *"Why do you think the temperature will be [insert student's numerical response] degrees?"* Then go to IV. ( ) Check off if student did not attempt this part of exercise.

b. offers any other response (or if no response is given), go to III.

III: **SAY:** *"Use the temperatures you recorded for the hot and cold water to predict the temperature of the mixture. If you want to do any figuring, you may use the paper provided. Put your prediction in* ***Part C*** *on your response sheet."* ( ) Check off if student did not attempt this part of exercise.

IV: **SAY:** *"Now pour the hot and cold water into the third cup and record the temperature in* ***Part D*** *on your response sheet."* ( ) Check off if student did not attempt this part of exercise.

ASSESSMENT CATEGORY	TITLE	SOURCE
Guided Investigation	Cold Pack	OAIP

COLD PACK

Professor Talensnik has helped to develop lab tests for high school chemistry in Ontario. He and colleagues constructed a set of chemistry tasks for the Ontario Assessment Instrument Pool (OAIP). Each task presents a problem for students to solve by applying chemical concepts and principles. Students are provided with standard lab equipment and a special lab kit for each specific investigation.

The first stage of the OAIP tasks requires that the students design the procedures that they believe will provide the data sufficient to solve the problem. After 15 minutes, the student designs are collected and evaluated. If the student plan is not complete, workable and safe, procedural steps will be given to the student before he/she is allowed to proceed.

The student's performance on the written plan and the report of the data collection and conclusions are evaluated by chemistry teachers. A total score of 35 points is possible for the chemistry lab test. Points are subtracted for each procedural step that had to be given before they began the experiment. Teachers are provided with detailed notes about material, procedures, and scoring.

COLD PACK

Procedural Step Model: Student Directions

I. **Introduction:** The medical "cold pack" operates on the principle that various salts absorb heat when they dissolve in water.

NOTE: You are permitted to use the Handbook of Chemistry and Physics, as well as your textbooks and any other materials provided by your teacher. You must leave ALL your "rough work" with the teacher.

II. **Special Lab Kit:** Unknown salt sample, thermometer, styrofoam cup, graduated cylinder (100 ml), balance, paper squares for use on the balance

Data: The heat absorbed per gram of substance dissolving (heat of solution) in joules per gram:

1.	ammonium chloride (NH_4Cl)	277 J/g
2.	ammonium perchlorate (NH_4ClO_4)	154 J/g
3.	potassium chloride (KCl)	231 J/g
4.	potassium nitrate (KNO_3)	346 J/g

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Guided Investigation	Cold Pack	OAIP

III. **Problem**: Using only the materials in the special lab kit, design an experiment to identify the unknown salt sample. The unknown is one of the salts listed in the "Data" on page 156.

The design must be written in detail on the **Student Sheet - Section A - Experimental Design**.

Do NOT proceed with the actual experiment work until the examiner has checked and approved the experimental design that you have suggested.

NOTE: You are also provided with a kit of "Standard Laboratory Glassware and Hardware."

IV. **Alternative Solution(s)**: Suggest, if you can, an alternate design for the experiment you have just written. The alternate design may involve the use of any equipment and supplies that you might suggest. The examiner will give you up to five additional points for any workable or reasonable alternative solution(s). Write the alternate solution(s) on the **Student Sheet - Section D - Alternate Solution(s)**.

V. **Timing**:

1. Time allowed for the experimental design - 15 minutes
2. Time allowed for the complete problem - 40 minutes

VI. **Test Format**: If you do not know how to proceed, you may request procedural steps from the teacher. The procedural steps are arranged in a sequential order. The teacher, after examining your progress, will give you the appropriate step(s). Each step will bring you closer to a complete procedure necessary to solve the problem. Each step that you request will result in a lower possible score that you can obtain.

After 15 minutes, your design(s) will be collected and evaluated. You will be given necessary procedural steps to follow unless you have presented a procedure which has been evaluated as complete and workable.

Procedural Steps:

Procedural Step No. 1: Find the mass of the salt sample using the balance.

Procedural Step No. 2: Measure 100 ml of water using the graduated cylinder. Pour the water into the styrofoam cup and measure its temperature.

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Guided Investigation	Cold Pack	OAIP

Procedural Step No. 3: Add the salt to the water in the styrofoam cup and stir until all the salt is dissolved; you may use the thermometer as a stirring rod if you stir very carefully.

Procedural Step No. 4:
1. Record the lowest temperature reached by the solution.
2. Calculate the quantity of heat absorbed by the salt-water system.

Procedural Step No. 5:
1. Divide the quantity of heat absorbed by the number of grams of salt used.
2. Compare this answer to the values provided under "data." The closest value identifies the unknown salt.

Suggested Procedures: There are many procedures for solving this problem. The set of five procedural steps outlines only one of the many possible procedures that might be designed by the students.

The teacher is responsible for approving and supplying the materials for any different procedure(s) that might be suggested. The teacher has the option to allow marks under the category of additional solution(s) for procedures other than "The Suggested Procedure."

Specific Laboratory and Security Procedures: The items listed in this section are in addition to the standard laboratory and security procedures that are in effect for each of the problems.

(a) KNO_3 and NH_4Cl give good results; they should be ground in a mortar to get a finer sample. This speeds up dissolving and reduces error.
(b) Provide each student with about 10g of the salt sample.
(c) Be careful that you do not use hydrated forms of the salts.
(d) The other salts are either hard to obtain or have low solubilities. They are added to the list as distracters.

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Guided Investigation	Cold Pack	OAIP

TEACHER'S SCORING GUIDE

Evaluation Scheme: (Total possible score: 35)

A. Design (10 points): (2 points per step)	Procedural Steps Required	B. Performance (10 marks): (2 points per step)
1. ________	1. ________	1. ________
2. ________	2. ________	2. ________
3. ________	3. ________	3. ________
4. ________	4. ________	4. ________
5. ________	5. ________	5. ________

Identity of the salt: ____________ (5 points)

C. Experimental error: (in calculating heat of solution)

	Pts.
> 50% ____________	0
30-50% ____________	1
10-30% ____________	3
< 10% ____________	5

D. Alternative procedure:

____

____

____

____

____

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Guided Investigation	Cold Pack	OAIP

Procedural Step Model Student Sheet

Student ______________________

Date ______________________

Score ______________________

A. Experimental Design	B. Observations

Identity of salt: ______________________

C. **Calculation of Experimental Error:**

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Guided Investigation	Cold Pack	OAIP

D. Alternate Solution(s) - Procedure

____

____

____

____

____

Notes and Calculations:

ASSESSMENT CATEGORY	TITLE	SOURCE
Practical Investigation	Acceleration	Rhode Island

ACCELERATION

This physics laboratory test was part of a Rhode Island program to recognize distinguished merit among their high school graduates. Students who wished to be so recognized responded to these tests after completing a paper-and-pencil achievement test in physics.

ACCELERATION

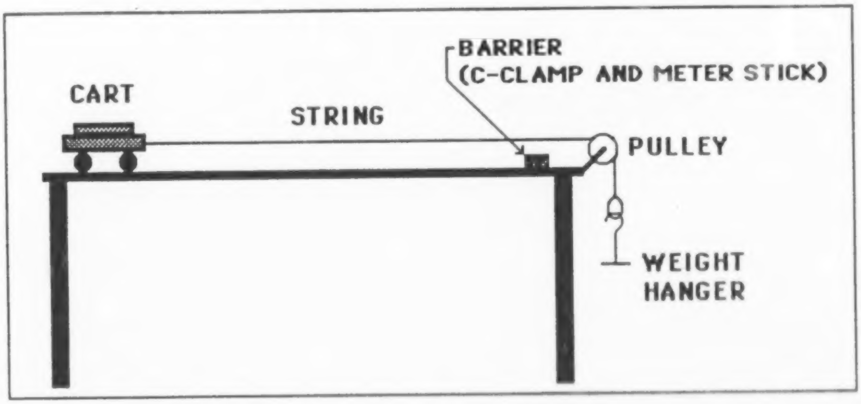
PHYSICS LABORATORY EXPERIMENT
NEWTON'S LAWS OF MOTION

Problem: To determine how acceleration varies when a constant force is applied to different masses. In this exercise, the same force will be applied to different masses and the resulting motion measured.

Materials: Cart, pulley, two meter sticks, balance, stopwatch, paper clips, string, mass set, C-clamp

Procedure:

1. Set up the equipment as shown in this diagram.



2. Suspend appropriate masses to the weight hanger until the frictional force acting on the cart is just offset. This happens when you can give the cart a very slight push and it moves at a constant speed across the table. DO NOT remove any of these masses at any time. (THE LENGTH OF THE STRING MUST BE ADJUSTED SO THE MASSES TOUCH THE FLOOR AS THE CART STRIKES THE BARRIER.) Record the total mass of the weight hanger and the masses at the bottom of the data table.

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Practical Investigation	Acceleration	Rhode Island

3. Measure the mass of your cart on a balance. Load the cart with two 1-kg masses and two 500 g masses (total 3 kg). Compute the total mass by adding 3 kg to the mass of the cart. Record the grand total as the Total Mass of the Cart on your data table.
4. Add 3 kg to the weight hanger. Have the cart move through the measured distance. Time the cart in traveling this measured distance. (Take the average of three trials.)
5. Keep the same suspended mass and weight hanger as part of the accelerating force for all trials. Remove 500 g from the cart and make a second timed trial.
6. Repeat the procedure described in step 5 for five more trials, removing 500 g from the cart before each trial. Your last trial should be made with the empty cart. You should have a total of seven timed trials.

Calculation: With your data and the information provided, calculate the acceleration (a) for each trial. Calculate the product of (m) and (a) for each trial.

Interpretation: What is the average value of the product of the total mass of the cart and weights and the acceleration of the cart? Is this consistent with Newton's Second Law of Motion? Explain what this value should be and why.

Useful Information: (Where $V_i = 0$)

$V_{avg} = S/t$
$V_f = 2\,V_{avg}$
$a = V_f/t$

On Earth, 1 kilogram weighs 9.8 Newtons. For the purposes of this experiment, consider the masses and the weight hanger to be the net external force on the string.

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Practical Investigation	Acceleration	Rhode Island

Note to Teacher: Prior to charging the students with performing this task, test the procedure to determine the appropriate accelerating force. A variety of masses should be given to students who will then "fine tune" the system.

DATA TABLE

TRIAL	TOTAL MASS (m) in kilograms	DISTANCE (S)	t sec	V_{avg} m/sec	V_f m/sec	a m/sec^2	ma in Newtons
1	3.0 + cart = _____						
2	2.5 + cart = _____						
3	2.0 + cart = _____						
4	1.5 + cart = _____						
5	1.0 + cart = _____						
6	0.5 + cart = _____						
7	0.0 + cart = _____						
						AVERAGE VALUE	

Mass of empty cart = ________ kg.

Masses and weight hanger = ________ kg.

Accelerating force (f) = ________ Newtons

ASSESSMENT CATEGORY	TITLE	SOURCE
Practical Investigation	Potato Cubes	Rhode Island

POTATO CUBES

This biology laboratory test was part of the same Rhode Island program as described for the "Acceleration" item previously mentioned. This test illustrates an activity in which students are focusing on a specific experimental design. Students are provided with detailed procedures to follow (with safety cautions highlighted) and suggestions for the analysis and interpretations of the findings.

POTATO CUBES

BIOLOGY LABORATORY EXPERIMENT
INFLUENCE OF TIME AND CONCENTRATION ON DIFFUSION

Problem: To determine the effect of time and concentration on the diffusion of potassium permanganate into potato cubes.

Materials: Potatoes; ruler; small beakers; 1%, 5%, and 10% potassium permanganate solutions; tweezers; scalpel; stopwatch

Procedure:

- With a scalpel, cut 13 cubes from a peeled potato. Each cube should measure 1 cm on each side.
- Fill three small beakers approximately 1/2 full with each of the three concentrations of potassium permanganate. Place four potato cubes into each beaker. Note the exact time the cubes were added to the solutions.
- With tweezers, remove one cube from the solution every five minutes.
- Slice each cube open with a scalpel. CAUTION: Slice away from fingers to avoid cuts. Carefully dry the scalpel before slicing each cube. Measure and record the average distance that the solution has diffused into each potato cube.

Analysis:

- Construct a data table showing your results.
- Construct a graph with the y-axis = the distance of diffusion, and the x-axis = time in solution. Be sure to label your graph clearly and show results with all three concentrations of solution.

Interpretation:

- State how the length of time in solution and the concentration of the solution influence the amount of diffusion.
- State how other factors that might influence diffusion.

ASSESSMENT CATEGORY	TITLE	SOURCE
Practical Investigation	Scoring Form	Rhode Island

SCORING FORM

The following is an example of a general scoring procedure used by Rhode Island in their "Distinguished Merit Program." The categories and topics can be used in a wide variety of laboratory related science tasks.

GENERAL SCORING GUIDE FOR SCIENCE LABORATORY TASKS

1. **Laboratory Skills**	**Quality Points**	**1**	**2**	**3**
● safety		☐	☐	☐
● massing and measuring techniques		☐	☐	☐
● proper handling of materials		☐	☐	☐
● proper handling of equipment		☐	☐	☐
● clean-up		☐	☐	☐

Possible points: 15
Awarded points: ____________

2. **Handling of Data**	**Quality Points**	**1**	**2**	**3**
● labeling of data		☐	☐	☐
● significant figures		☐	☐	☐
● accuracy		☐	☐	☐
● units		☐	☐	☐
● utilization of data		☐	☐	☐

Possible points: 15
Awarded points: ____________

3. **Interpretations**	**Quality Points**	**1**	**2**	**3**
● title of graph		☐	☐	☐
● nature of investigation and purpose		☐	☐	☐
● data/calculation		☐	☐	☐
● conclusion		☐	☐	☐
● questions/further investigation		☐	☐	☐

Possible points. 15
Awarded points: ____________

Student: ____
Observer: ____
Date: ____

TOTAL POINTS: ____________/45

ASSESSMENT CATEGORY	TITLE	SOURCE
Practical Investigation	Acid/Base	LABS

ACID/BASE

This chemistry lab test was part of a monograph entitled Laboratory Assessment Builds Success (LABS). This publication was prepared by the participants of the Institute for Chemical Education at the University of California at Berkeley from 1987-89 under the direction of Marjorie Gardner. The monograph presents 18 key labs for high school chemistry with sections designed for the students and sections for the teachers. Several methods of assessment were recommended for use with these activities, including lab practical tests. The following lab practical test on the acid-base properties of oxides was from LABS.

ACID/BASE

Laboratory Practical: Acid-base properties of oxides

Notes to Teacher:

Safety:

- Protective glasses and aprons must be worn.
- Solutions are corrosive and should be handled with care.

Materials:

- Non-consumables: one glass stirring rod, six test tubes (13 x 100 mm)
- Consumables: methyl orange indicator; 0.1 M KOH (0.6 g KOH per 100 ml soln.); 0.1 M NaOH (0.4 g NaOH per 100 ml soln.); 0.1 M HNO_3 (0.5 g concentrated HNO_3 per 100 ml soln.); 0.1 M H_2SO_4 (0.6 g concentrated H_2SO_4 per 100 ml soln.)

Advanced Preparation:

The acid and other solutions should be placed in dropper bottles labeled with letters or numbers. Other acids and bases could be used. Depending on their level, students might need written directions for setting up this investigation.

Add about 1 ml (20 drops) of each solution to separate test tubes and add one drop of methyl orange indicator to each tube. The procedure would still work if you add one drop of each solution to 1 ml of water and add the methyl orange to this. The results should be good in either case.

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Practical Investigation	Acid/Base	LABS

STUDENT WORKSHEET

ACID/BASE

Directions to Students:

CAUTION: Solutions are corrosive. Wear protective glasses and aprons during this investigation. You have been given four dropper bottles, each containing an oxide dissolved in water. You also have an indicator which is red in a very acidic solution, orange in a mildly acidic solution, and yellow in a basic solution.

Write a procedure and use it to determine which dropper bottles contain metallic oxides and which contain nonmetallic oxides.

Procedure:

Observations:

Solution	Color	Interpretation			
		A	B	M	N
1.					
2.					
3.					
4.					

A = Acid B = Base M = Metal N = Non-metal

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Practical Investigation	Acid/Base	LABS

Anticipated Student Results: The following observations are typical.

Solution	Color	Interpretation
1. KOH	yellow	basic, metal oxide
2. NaOH	yellow	basic, metal oxide
3. HNO_3	red	acidic, non-metal oxide
4. H_2SO_4	red	acidic, non-metal oxide

The following is a chart for assessing the skills performed in the investigation:

STUDENT NAME	EXPERIMENT DESIGN	CAREFUL OBSERVATIONS (color)	INTERPRETATIONS BASED ON OBSERVATIONS (score 0-8 points)

ASSESSMENT CATEGORY	TITLE	SOURCE
Practical Investigation	Hooke's Law	NORC-UB

HOOKE'S LAW

In the Development of Twelfth Grade Science Assessment Project (involving the National Opinion Research Center and the State University of New York at Buffalo), Bock and Doran created a set of science laboratory tests for the high school courses in biology, chemistry, and physics. These tests are divided into two parts. In Part A (30 minutes), the students are requested to design an experiment to solve a stated problem, applying specific science concepts and laboratory procedures. They are expected to state a hypothesis, describe a general strategy or plan, list a series of detailed procedures and prepare a method of organizing the observations/data that are to be collected. After 30 minutes, their responses to Part A are collected. To eliminate "double jeopardy" and the difficulty of instantly evaluating a number of students' plans, Bock and Doran provide a detailed set of procedures for the students to follow in Part B. Students are expected to manipulate apparatus, collect and organize data or observations, perform calculations, graph results, and formulate conclusions.

The students' responses to both Parts A and B are scored by experienced science teachers who have been trained with the scoring procedures. The scoring form shown below will provide a compact summary of their performance. Descriptions of the criteria for each point value for each skill have been developed to guide the scoring. Thirdly, a collection of student answers are assembled to illustrate for the scorers what to expect for each point value for each laboratory skill. This scoring procedure is a modification of that used successfully in Israel by Tamir. The item illustrated here is from the physics set. A set of six tasks have been prepared for each science area. These tasks were written to use terminology, procedures, and equipment that students had experienced during their science class. The items include problems and tasks that are somewhat novel, and not direct replication of familiar laboratory activities. For instance, in item 1.2, the mass of the object is greater than the maximum capacity of either scale used individually. In task 1.3, students are not provided masses that weigh 15 Newtons. As part of this illustration, we have included the directions for administrating the test, the students' booklet, the scoring form, and the scoring criteria.

HOOKE'S LAW

Instructions to Students:

As students enter the room, they should be instructed to sit at a station. When the students are seated, let them look around the room, then instruct them to look at the equipment in front of them and find their test booklet. Tell the students NOT to touch the equipment until they are told to do so.

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Practical Investigation	Hooke's Law	NORC-UB

When the students have settled, the supervising teacher should read the following text. For ease in administration, all sections to be read are printed in italics.

Good morning (afternoon). My name is ____________________. Today, we are going to perform some science experiments. A booklet explaining the experiment you are to do should be in front of you. If you cannot find your booklet or you do not have a pencil, please raise your hand.

DISTRIBUTE NEEDED TEST BOOKLETS AND/OR PENCILS.

Now that each of you has a test booklet and pencil, we are ready to begin. Listen very carefully to the instructions. Please write your name, your sex, your school, and today's date in the space provided on the front page of the booklet.

GIVE STUDENTS TIME TO DO THIS.

This is a science laboratory test. The things that you will need have been set out in front of you. Your booklet will give you specific directions for this experiment. Write the results of your experiments in your booklet.

The science laboratory test is very different than the other science tests that you completed. We are trying to see how testing of science skills can be done in U.S. high schools. Your responses are very important to us in this project.

Your individual performance will not be used as part of any science grade or evaluation for you. We ask you to do your best so we can learn as much as possible about this alternative method of testing.

MAKE SURE EVERYONE IS AT THE CORRECT PAGE IN THE BOOKLET AND THAT THEY HAVE ALL THE MATERIALS THEY NEED.

The physics test is organized into two parts, A and B. In Part A, you will do three tasks. At each of the first two tasks, you should spend five minutes. The third task has been scheduled for twenty minutes of your time. When the time allowed for Part A (30 minutes) has elapsed, I will collect your booklet for Part A and give you the booklet for Part B.

Once you have started your experiment, I cannot help you. I can only assist you right now if you have any materials missing.

We are now ready to begin the test. Are there any questions? Do your best. You may begin.

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Practical Investigation	Hooke's Law	NORC-UB

AFTER 30 MINUTES, COLLECT THE PART A BOOKLETS AND GIVE THE PART B BOOKLETS.

You will have 50 minutes for Part B. You may begin.

AFTER THE STUDENTS HAVE COMPLETED THE TEST, READ THE FOLLOWING INSTRUCTIONS.

This is the end of the science laboratory test. Please pass in your booklet and pencils.

Thank you for being so attentive and cooperative during the test. Please wait to be dismissed.

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Practical Investigation	Hooke's Law	NORC-UB

SCIENCE LABORATORY TEST

PHYSICS

Task #1
Hooke's Law

NAME ______________________________ SEX ______________

SCHOOL ______________________________ DATE ______________

PART A
30 minutes

These tests are being developed through the University of Buffalo
with support from the National Science Foundation.
Copyright 1990

Introduction:

In this laboratory test, you will be asked to do three different activities.

ACTIVITY	INSTRUCTIONS
1.1 1.2	You have five minutes to complete your work on each activity. Follow the directions on each activity and record all your work on these answer sheets.
1.3	You will be given a problem and a list of materials. You will have 20 minutes to design a strategy for solving the problem. Please record all your answers on these sheets.

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Practical Investigation	Hooke's Law	NORC-UB

LABORATORY TEST
ITEM #1.1

DRAWING A VECTOR DIAGRAM

Problem:

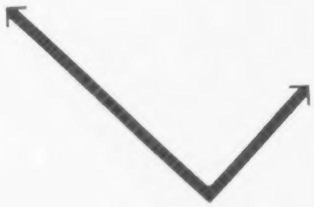
Refer to the scale diagram of two force vectors as shown below.

1. Draw the resultant of the two forces. (Be certain to show its direction.)
2. Determine and record its magnitude and units.

Materials: Ruler, protractor

Diagram

1 cm = 1 Newton



Ans. ______________

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Practical Investigation	Hooke's Law	NORC-UB

LABORATORY TEST
ITEM #1.2

MEASURING A WEIGHT USING TWO SPRING SCALES

Problem:

Design and implement a method to measure the weight (in Newtons) of object A.

1. Create a labeled diagram of the positions of the scales and object A.
2. Record your measurements.

Materials: Object A, two spring scales, thread

Ans. Mass of Object A ________________

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Practical Investigation	Hooke's Law	NORC-UB

LABORATORY TEST
ITEM #1.3

Introduction:

This laboratory test presents a problem and lists materials. You will have 20 minutes to plan and design an experiment to solve the problem.

Problem:

You will remember that Hooke's Law describes the relationship between the force applied to a spring and its elongation (stretch). That is, if a force stretches a spring, the elongation is directly proportional to the force applied. $F = kx$. Where: F = force in Newtons, k = spring constant in Newtons/meter, x = elongation in meters.

Imagine that you work for a spring manufacturer and your job is to determine whether Spring A will elongate to exactly 0.3 meters when a force of 15 Newtons is applied. Your job is to conduct an experiment which will determine whether Spring A is capable of this task.

a) Under the heading "Procedure," list in order the steps of the procedure you will use to solve the problem. Include a diagram to help illustrate your plans for the experiment. Include any safety procedures you would follow.
b) Construct a data table or indicate any other method that you could use to record the observations and results that will be obtained.
c) At the end of 20 minutes, your answer sheet will be collected.

Please note: In Part 1, you are not to proceed with any part of the actual experiment. You are just to plan and organize a way to investigate the problem.

Materials: A set of masses which totals 1 kg, Hooke's Law Apparatus, Spring A, C-clamp, graph paper, calculator

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Practical Investigation	Hooke's Law	NORC-UB

EXPERIMENT DESIGN

Organize your experiment under the following headings: "Procedure" and "Result/Observations." Use the front and back of these sheets if necessary.

Procedure:

Results/Observations (e.g., data table, graphs)

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Practical Investigation	Hooke's Law	NORC-UB

PART B - 50 MINUTES

Introduction:

You will have 50 minutes to complete this part. You have been provided with a detailed procedure (see next page) which you are to follow. Record your work for Part 2 on the answer sheet under the appropriate headings.

Problem:

You will remember that Hooke's Law describes the relationship between the force applied to a spring and its elongation (stretch). That is, if a force stretches a spring, the elongation is directly proportional to the force applied. $F = kx$. Where: F = force in Newtons, k = spring constant in Newtons/meter, x = elongation in meters.

Imagine that you work for a spring manufacturer and your job is to determine whether Spring A will elongate to exactly 0.3 meters when a force of 15 Newtons is applied. Your job is to conduct an experiment which will determine whether Spring A is capable of this task.

a) Perform the experiment by following the steps outlined in the procedure.
b) Under the heading "Data," record the results of the experiment. Use statements, descriptive paragraphs, and tables of data where appropriate.
c) Under the heading "Calculations," show all equations and calculations used.
d) Under the heading "Conclusion," provide graphs and give an interpretation of your results. What did you learn from the experiment?
e) At the end of the 50 minutes, your answer sheets will be collected.

Materials: A set of masses which totals 1 kg, Hooke's Law Apparatus, Spring A, C-clamp, graph paper, calculator

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Practical Investigation	Hooke's Law	NORC-UB

EXPERIMENT REPORT #1 - PART B

Complete the procedure as given on these sheets.
Record your data in the table provided.

Procedure:

1. Check the attachment of Spring A and adjust the pointer to zero on the scale.
2. Attach increasing amounts of mass. Record both the total mass and the total elongation of the spring of each trial. Use the data table provided.
3. Calculate the force of gravity acting on each mass in Newtons. ($W = mg$, $g = 9.8 m/sec.^2$)
4. Plot a graph of force against elongation, where the scale for values of force include 0 to 20 Newtons and values for meters include 0 to 0.4 meters.
5. Use the graph to determine the force constant of the spring. **Show all calculations.** The slope of the line is the force constant k.

$$k = \frac{\triangle F}{\triangle x}$$

where: k = force constant
$\triangle F$ = force
$\triangle x$ = elongation

6. Determine whether Spring A can stretch to exactly 0.3 meters with a force of 15 Newtons applied, by extending the graph through to force = 15 Newtons.
7. Report your analysis of this problem under "Conclusion."

Data:

TRIAL #	MASS (kg) (1,000 g = 1 kg)	FORCE (N) W = mg	ELONGATION WITH LOAD (m)
1			
2			
3			
4			
5			
6			
7			
8			

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Practical Investigation	Hooke's Law	NORC-UB

EXPERIMENT REPORT #1 - PART B

Record your calculations and conclusion on these sheets.

Calculations:

Conclusion:

Based on a graphical analysis of the data you collected in this experiment, discuss your conclusion as to whether Spring A could stretch to exactly 0.3 meters if a force of 15 Newtons is applied. Be certain to explain how you used your data to arrive at this conclusion.

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Practical Investigation	Hooke's Law	NORC-UB

SCORING FORM - SCIENCE LABORATORY TEST

SCHOOL AND STUDENT I.D. NO.: ________________ DATE: __________

READER I.D. NO.: ____________________________ TIME: ____:____

SUBJECT AREA: G B C P TASK: 1 2 3 4 5 6

Points:

	NR	0	1	2	NA		**Part A: Experiment Design**
1.	☐	☐	☐	☐	☐	1.	Statement of hypothesis
2.						2.	Procedure for investigation
	☐	☐	☐	☐	☐		- Description of a strategy or plan
	☐	☐	☐		☐		- Diagram of equipment
	☐	☐	☐		☐		- Safety procedures
3.	☐	☐	☐	☐	☐	3.	Plan for data collection

Points:

	NR	0	1	2	3	NA		**Part B: Experiment Report**
4.	☐	☐	☐	☐	☐	☐	4.	Quality of the observations/data
5.							5.	Graph
	☐	☐	☐	☐		☐		- Has an appropriate title
	☐	☐	☐	☐		☐		- Axes labeled with correct variables/units
	☐	☐	☐			☐		- Placement of dependent variable
	☐	☐	☐	☐		☐		- Scale of the graph is appropriate
	☐	☐	☐	☐		☐		- Points are plotted accurately
	☐	☐	☐	☐		☐		- Curve is appropriate for the data trend
6.							6.	Calculations
	☐	☐	☐			☐		- Correct use of units
	☐	☐	☐			☐		- Accuracy of calculations
	☐	☐	☐	☐		☐		- Knowledge of relationships
7.							7.	Forms a conclusion from the experiment
	☐	☐	☐	☐		☐		- Consistent with data collected
	☐	☐	☐	☐		☐		- Consistent with scientific theory/ principle

Reader Comments: ____

____

____

NR = No Response NA = Not Applicable

NOTE: Use a No. 2 pencil or erasable ball-point pen.
Print numbers in block style: 0 1 2 3 4 5 6 7 8 9
Mark rating boxes with an "X."
Erase all errors completely.

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Practical Investigation	Hooke's Law	NORC-UB

CRITERIA FOR SCORING FORM
SCIENCE LABORATORY TEST

PART A: EXPERIMENT DESIGN

1. Statement of Hypothesis	Points
• Adequate and complete	2
• Partial understanding	1
• No understanding of the problem	0
• No answer	NR
2. Procedure for Investigation	
• Description of a strategy or plan	
-- Complete and reasonable	2
-- Partial description	1
-- Inappropriate plan	0
-- No plan	NR
• Diagram of experiment	
-- Appropriate diagram	1
-- Inappropriate diagram	0
-- No drawing is presented	NR
• Safety procedures	
-- Appropriate safety precautions	1
-- Inappropriate safety precautions	0
-- No safety precautions mentioned	NR
3. Plan for Data Collection	
• Complete and reasonable	2
• Partial description	1
• Inappropriate plan	0
• No plan	NR

PART B: EXPERIMENT REPORT

4. Quality of the Observations/Data	
• Complete and consistent	3
• Incomplete but consistent	2
• Complete but inconsistent	1
• Incomplete and inconsistent	0
• No observations/data are reported	NR

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Practical Investigation	Hooke's Law	NORC-UB

5. Graph

● Has an appropriate title	
-- Completely correct title	2
-- Partially correct title	1
-- Incorrect title	0
-- No title	NR
● Axes labeled with correct variables/units	
-- Correct variables and units on both axes	2
-- Partially correct labeling of axes	1
-- Incorrect variables and units	0
-- No variables or units labeled	NR
● Placement of dependent variable	
-- Dependent variable is plotted on the vertical axis	1
-- Correct placement of dependent variable	0
-- Dependent variable not labeled	NR
● The scale of the graph is appropriate	
-- Correct and proper scale on both axes	2
-- Correct and proper scale on one axis, not both	1
-- Both scales inappropriate	0
-- Scales not presented	NR
● Points are plotted accurately	
-- Points correctly plotted	2
-- Points partially correct	1
-- Points inaccurately plotted	0
-- No data points plotted	NR
● The curve is appropriate for the data trend	
-- The curve is correctly drawn	2
-- The curve is slightly inaccurate	1
-- The curve is very inaccurate	0
-- No curve is drawn	NR

6. Calculations

● Correct use of units	
-- Calculations include correct units	1
-- Incorrect units used	0
-- No units shown	NR
● Accuracy of calculations	
-- Consistent with the data	1
-- Inconsistent with the data	0
-- No calculations performed	NR
● Knowledge of relationships	
-- Complete understanding of relationships	2
-- Partial understanding	1
-- Inaccurate or incorrect understanding	0
-- No relationship described	NR

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Practical Investigation	Hooke's Law	NORC-UB

7. Forms a Conclusion From the Experiment

- Consistent with data collected	
-- Correct and properly justified	2
-- Partially correct, not fully justified	1
-- Incorrect conclusions	0
-- No conclusions are given	NR
- Consistent with scientific theory/principle	
-- Correct and properly justified	2
-- Partially correct, not fully justified	1
-- Incorrect conclusions	0
-- No conclusions are given	NR

ASSESSMENT CATEGORY	TITLE	SOURCE
Practical Investigation	Identifying Gases	IPS

IDENTIFYING GASES

Several laboratory tests were prepared for use with the Introductory Physical Science (IPS) program. These tests were designed to assess the students' ability to collect data and formulate conclusions. In the test included here--Collecting and Identifying Gases--the students are expected to perform several tests which they have learned in the IPS course. Extensive notes are provided for the teacher, including a summary of possible student responses. Please notice the specific notes about safety.

COLLECTING AND IDENTIFYING GASES

Introductory Physical Science
Laboratory Test No. 1

To the Student:

Arrange apparatus to produce and collect a gas. Two of the substances you have been given will produce a gas when dropped into water. The third substance is a liquid which must be heated to evolve gas.

Collect samples of the gas produced by each of the substances. Find out as much as you can about the properties of the gas produced in each case, using the tests with which you are now familiar.

Describe each gas as carefully and as completely as you can.

Identify the gas if you can.

Apparatus and Materials:

Peg board; two clamps; six test tubes; one one-hole No. 2 stopper; four no-hole No. 2 stoppers; one right-angle glass bend; rubber tubing; one plastic bucket or large pan; one burner; wood splints; matches; lime water (45 ml); water; paper toweling; three unknown substances labeled "X," "Y," and "Z."

Teacher's Notes:

This test will take one to two periods. The chemicals used can be listed as "X" (Alka-Seltzer tablet), "Y" (calcium carbide, 1.2 g), and "Z" (solution of ammonium chloride and sodium nitrite, 50 ml). You may wish to hand out printed instructions for mixing before the experiment is performed. **Be sure your students wear safety glasses and aprons!**

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Practical Investigation	Identifying Gases	IPS

Have your students use one Alka-Seltzer tablet (chemical "X") in a quarter of a test tube of water.

Calcium carbide (chemical "Y") reacts very rapidly. No more than three beads of calcium carbide should be added to one half of a test tube of water at one time. If more gas is needed later, another three beads can be added.

CAUTION: Do not let students use excessive amounts of calcium carbide, since the gas produced is explosive.

You can prepare the liquid sample (chemical "Z") in advance by dissolving 10.0 g of sodium nitrite **(not sodium nitrate!)** and 10.0 g of ammonium chloride in 150 ml of water. (CAUTION: The two solids must be heated together only in water. If heated dry, they may explode!) It is best to have some extra solution on hand. Have your students use only one quarter of a test tube of the liquid when they produce the gas. Once the reaction starts, gas is evolved quite rapidly. Warn students to be ready to change collecting tubes quickly.

As in similar previous experiments, the first test tube of each gas should be discarded.

The limewater and flammability tests are straightforward, but conclusions about the density will require careful observation of the flammability tests for acetylene and nitrogen with upright and inverted test tubes respectively.

Sample Student Results:

The gas produced when "X" is added to water extinguishes a glowing split, is more dense than air, and turns limewater milky. It is, therefore, carbon dioxide.

The gas produced when "Y" is added to water burns with a yellow, sooty flame, is slightly less dense than air, and does not affect limewater. It is not hydrogen, carbon dioxide, or air.

The gas produced when "Z" is heated extinguishes a glowing splint and flaming splint. It has about the same density as air and does not affect limewater. It is not hydrogen, carbon dioxide, the gas produced by the gray solid, or air.

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Practical Investigation	Identifying Gases	IPS

A Summary of the Student's Results:

SUBSTANCES	FLAMMABILITY				LIMEWATER TEST
	MOUTH UP (60 SEC.)		MOUTH DOWN (60 SEC.)		
	GLOWING SPLINT	FLAMING SPLINT	GLOWING SPLINT	FLAMING SPLINT	
X	goes out	goes out	glows	goes out	cloudy
Y	glows	goes out	gas burns	gas burns	clear
Z	goes out	goes out	goes out	goes out	clear

Conclusions:

a. I conclude that the gas produced by substance X is carbon dioxide.

b. I cannot identify the other gases.

ASSESSMENT CATEGORY	TITLE	SOURCE
Practical Investigation	Photosynthesis	BSCS

PHOTOSYNTHESIS

Shortly after the development of the BSCS instructional material, several efforts were undertaken to produce innovative student assessment instruments. The Process of Science Test (a multiple choice test) and other multiple choice achievement tests of the BSCS content were developed and distributed commercially through the Psychological Corporation. Consistent with the inquiry approach of the BSCS program, tests were developed to assess students' laboratory skills. In a BSCS newsletter, Tamir and Glassman described the Laboratory Tests for BSCS Students, which included a set of six lab tests on the following topics: Photosynthesis; Daphnia - Alternation of Activity; Measuring the Rate of Human Respiration; Grasshopper Respiration; Yeast Fermentation; and Water Relationships of a Plant Tissue.

These laboratory tests required one to two class periods for students to complete. The students were presented with an experimental set-up and a series of 10 to 15 questions. These questions probed a variety of laboratory skills ranging from the planning and design of experiments, the manipulation of apparatus to collect data, and the formulation of conclusions from the findings of the investigation.

The scoring of the performance tasks was conducted by experienced biology teachers. They used a set of criteria established for each task with the student written responses. Tamir has extended this biology lab testing to the Israeli High School Matriculation Exam in Biology.

PHOTOSYNTHESIS

Problem 1. Measuring the Rate of Photosynthesis

Set-Up:

On the table are three beakers filled with water. In each is an inverted funnel containing several sprigs of fresh elodea. On the funnels are calibrated test tubes. The first set-up is in the direct light provided by a 100-watt lamp. The second set-up is about one meter distance from the lamp. The third set-up is completely concealed under a heavy paper cylinder. There are also two liter bottles containing a solution of $NaHCO_3$. (If the set-ups are arranged about one hour before the students arrive, there is a clearly discernible difference in gas level in the test tubes.)

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Practical Investigation	Photosynthesis	BSCS

TO THE STUDENT: DIRECTIONS, PROCEDURES, AND QUESTIONS	EXAMINER'S GUIDE: TESTING PROCEDURES AND EVALUATION OF RESPONSES
1. Examine the rate of photosynthesis of the three set-ups in front of you. Write the results of your observations.	1. The student will have to measure the rate of photosynthesis by observing the accumulation of gas in each of the three test tubes for 10 minutes.
2. What is the control in this experiment?	2. This experiment has no control. If the student did not mention this, he would lose five points.
3. How would you explain the results? Indicate the major processes occurring in each of the set-ups.	3. The student's explanation should include an indication of the processes going on and the reasons for the observed differences.
4. What is the gas that collects in each of the test tubes? How can you test this?	4. The student will suggest how the test is to be carried out--but the student is not asked to perform the actual test.
5. Why did we use a water plant (elodea) in this experiment?	5. The dependence of the method of measurement on the type of plant selected is to be explained.
6. You have more elodea twigs. Design an experiment whereby you will be able to measure the rate of photosynthesis in a different way. Before you begin, call the examiner and explain your plan. Perform the experiment and write the results.	6. If the student had not proposed to count the bubbles, the examiner will propose this method and subtract five points.
7. How would you speed up the rate of photosynthesis in the experiment you designed? (HINT: Look at the equipment and materials on the table.)	7. Here the student may use light or add $NaHCO_3$.

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Practical Investigation	Photosynthesis	BSCS

TO THE STUDENT: DIRECTIONS, PROCEDURES, AND QUESTIONS	EXAMINER'S GUIDE: TESTING PROCEDURES AND EVALUATION OF RESPONSES
8. Determine quantitatively if, and to what extent, you succeeded in speeding up the rate of photosynthesis. Write your results. Is there any control in this experiment?	8. The student is expected to determine the rate of photosynthesis before and after giving the treatments. In order to change light intensity, s/he can put the plant at a certain distance from the lamp and gradually bring it closer. In the case of $NaHCO_3$, s/he can add a small amount, measure, than add some more, etc. In this experiment, the first measurement serves as a basis for comparison (like control) and, therefore, the procedure is legitimate.
9. What are your conclusions from all the experiments that you did?	9. The conclusions will relate to the different methods of measurement and to various external factors which may change the rate of photosynthesis.
10. Write down a hypothesis based on the results of the experiments you performed.	10. An example of a hypothesis: "The rate of photosynthesis increases as light intensity increases."
11. Describe, in short, how you can test your hypothesis experimentally.	11. Experiment: We may measure the rate of photosynthesis using two lamps similar in intensity to the original one, then three lamps, etc.

ASSESSMENT CATEGORY	TITLE	SOURCE
Extended Investigation	Cost of a Shower	Connecticut

COST OF A SHOWER

The extended investigation, as a formal assessment tool, has been championed by Baron and colleagues. Much of the development and trial testing of these tasks have been done by science teachers in Connecticut and other cooperating states. This form of assessment occurs across several stages and several class periods. At some stages, students work individually, and in other stages, they work in small groups.

As one might expect, there are several sets of student material for these tasks, as well as several sets of guidelines for evaluation.

WHAT DOES IT COST TO TAKE A SHOWER?

In Your Group:

1. Brainstorm a list of all of the energy transformations that are involved in taking a shower. Be careful to consider everything that is required to take a shower.
2. Design a study to investigate the cost of all of these energy transformations. Include the following in this design:
   a. How you will go about finding all of the energy costs.
   b. All of the data that you need to collect.
   c. Any equipment you might need.
   d. How you will reach your conclusion.
   e. Any safety considerations that you might need to follow.
3. Submit your design to your teacher for approval.
4. Carry out your study.
5. Present your results to the class, along with some form of visual aid to help make your presentation clearer. Include a discussion of how accurate you think your final results are. Discuss possible sources of error and explain how they affected your results.

The scoring criteria for this activity are provided on the Performance Task Evaluation Sheet and the Participation in Discussion Self-Assessment Questionnaire which follow.

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Extended Investigation	Cost of a Shower	Connecticut

PERFORMANCE TASK EVALUATION SHEET

Performance Task Title: *What Does It Cost to Take a Shower?*

Student's Name: ____________________

Class: ____________________ **Date:** ____________________

Teacher's Name ____________________

PERFORMANCE GOALS

THE STUDENT CAN...	EXCEEDS GOAL	MEETS GOAL	NEARS GOAL	NOT YET
1. Apply the basic concepts of energy transformations to solve a problem.				
2. Identify appropriate information and steps needed to solve a problem.				
3. Make and support reasonable estimates based on appropriate assumptions.				
4. Assess and synthesize appropriate information found in various sources.				
5. Perform necessary procedures, calculations, and analyses to reach a reasonable conclusion.				
6. Identify sources of error and explain how they affected the results.				
7. Communicate orally the strategy and outcomes of the study using appropriate visual aids.				

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Extended Investigation	Cost of a Shower	Connecticut

PARTICIPATION IN DISCUSSION SELF-ASSESSMENT QUESTIONNAIRE

NAME ______________________________ DATE ______________

Circle Yes or No

A. Group Participation

1. YES NO Participated in group discussion without prompting
2. YES NO Did my fair share of the work
3. YES NO Tried to dominate the group; interrupted others; spoke too much

	Almost Always	Often	Sometimes	Rarely
Participated in the Group's Activities	○	○	○	○

B. Staying on Topic

1. YES NO Paid attention; listened to what was being said and done
2. YES NO Made comments aimed at getting the group back to the topic
3. YES NO Got off the topic or changed the subject

	Almost Always	Often	Sometimes	Rarely
Stayed on Topic	○	○	○	○

C. Offering Useful Ideas

1. YES NO Gave ideas and suggestions that helped the group
2. YES NO Offered helpful criticism and comments
3. YES NO Influenced the group's decisions and plans

	Almost Always	Often	Sometimes	Rarely
Offered Useful Ideas	○	○	○	○

D. Consideration

1. YES NO Made positive, encouraging remarks about group members and their ideas
2. YES NO Gave recognition and credit to others for their ideas
3. YES NO Made inconsiderate or hostile comments about a group member

	Almost Always	Often	Sometimes	Rarely
Was Considerate of Others	○	○	○	○

(continued)

ASSESSMENT CATEGORY	TITLE	SOURCE
Extended Investigation	Cost of a Shower	Connecticut

Circle Yes or No

E. **Involving Others**
  1. YES NO Got others involved by asking questions, requesting input, or challenging others
  2. YES NO Tried to get the group working together to reach group agreements
  3. YES NO Seriously considered the ideas of others

	Almost Always	Often	Sometimes	Rarely
Involved Others	○	○	○	○

F. **Communicating**
  1. YES NO Spoke clearly; was easy to hear and understand
  2. YES NO Expressed ideas clearly and effectively

	Almost Always	Often	Sometimes	Rarely
Communicated Clearly	○	○	○	○

After filling out this form, on a separate sheet of paper, please elaborate upon your strengths and the things that you need to work on during class discussion.

SHORT INDIVIDUAL PERFORMANCE TASK

On Your Own:

Now that you understand all of the energy costs that go into taking a shower, design the least expensive shower that you can imagine. Include all of the energy transformations that are used in this new shower.